





```

QY 1561 GTGGCTGTTGGACCCCGGCTGCTGAGGAGACACTGTCTACGAGCCCTTGTGAG 1620
    |||||||
Db 1561 GTGGCTGTTGGACCCCGGCTGCTGAGGAGACACTGTCTACGAGCCCTTGTGAG 1620
QY 1621 CATTCCGGAAGATTTTTCGACTACAGTTTATGATATTCAGATGTCAGTTCACAT 1680
    |||||||
Db 1621 CATTCCGGAAGATTTTTCGACTACAGTTTATGATATTCAGATGTCAGTTCACAT 1680
QY 1661 GATGACACAAATCTCATCTGGGACTTCTTAATATGATCCAGCTGCCAAGTGAACCCCCC 1740
    |||||||
Db 1661 GATGACACAAATCTCATCTGGGACTTCTTAATATGATCCAGCTGCCAAGTGAACCCCCC 1740
QY 1741 CGTTCCTTCTGCAACATACCTACATCTCCGATTAATATACCATACACTGACTCAT 1800
    |||||||
Db 1741 CGTTCCTTCTGCAACATACCTACATCTCCGATTAATATACCATACACTGACTCAT 1800
QY 1801 ACTTGCCAGAGACCCATTAAGTTGGGATTTTAAAGTATGCAATGCAATACAGATGAGC 1860
    |||||||
Db 1801 ACTTGCCAGAGACCCATTAAGTTGGGATTTTAAAGTATGCAATGCAATACAGATGAGC 1860
QY 1861 AACCAAGTAACTAACTAACTACTGCTCCAGTTTCCCTGGAGTACCGAGAGAGAGGCTT 1920
    |||||||
Db 1861 AACCAAGTAACTAACTAACTACTGCTCCAGTTTCCCTGGAGTACCGAGAGAGAGGCTT 1920
QY 1921 TGAGACTGCTGTTGGAGACAGTGTGCTGACGTGCGCCGAGACGGTCTACTAGACACA 1980
    |||||||
Db 1921 TGAGACTGCTGTTGGAGACAGTGTGCTGACGTGCGCCGAGACGGTCTACTAGACACA 1980
QY 1981 ACTGACTGCTGCTGAGTGTGCTGATGAGAGATGCTTCTGATCAATGGAATGATTTGGAAC 2040
    |||||||
Db 1981 ACTGACTGCTGCTGAGTGTGCTGATGAGAGATGCTTCTGATCAATGGAATGATTTGGAAC 2040
QY 2041 TTTTAAACCTCCCTCCTCTCTCTCTTTCACCTGTCACCTAGTTTTCCTCATTTGGTTC 2100
    |||||||
Db 2041 TTTTAAACCTCCCTCCTCTCTCTCTTTCACCTGTCACCTAGTTTTCCTCATTTGGTTC 2100
QY 2101 CAGACAAAGTGACTATTAATATATTTAGTGTGTCAGCAAAAAA 2151
    |||||||
Db 2101 CAGACAAAGTGACTATTAATATATTTAGTGTGTCAGCAAAAAA 2151

RESULT 2
ID AAA51229 standard; DNA: 2151 BP.
AC AAA51229;
XX 26-SEP-2000 (first entry)
DT Human beta-Trcp coding sequence.
DE
XX E3 ubiquitin ligase; beta-Trcp; F-box; WD protein; I-kappa-B; inhibitor;
KW nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;
KW anti-inflammatory; immunosuppressive; cytosolic; ds.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH 70..1779
FT CDS /*tag= a
FT /product= Human_Trpc
XX
XX WO200034447-A2.
XX
XX 15-JUN-2000.
XX
XX 10-DEC-1999; 99WO-US29371.
XX
XX 10-DEC-1998; 98US-021060.
XX
XX (SIGN-) SIGNAL PHARM INC.
XX (YISS) YISSUM RES & DEV CO.

```

```

XX
PI Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzudai A;
PI Lavon I, Yaron A;
XX
DR MPI: 2000-431294/37.
DR P-PDB: AAY6697.
XX
PT Polypeptide enhancing phosphorylated Ikkappa ubiquitination useful for
PT treating disorder associated with NF-kappaB activation e.g. cancer,
PT comprising amino acid sequence of human E3 ubiquitin ligase or its
PT variant
XX
PS Disclosure: Page 72; 77pp; English.
XX
CC This DNA encodes human beta-Trcp, an F-box/WD protein family member,
CC which has been shown to have homology to human E3 ubiquitin ligase (E3).
CC E3 enhances ubiquitination of phosphorylated I-kappa-B, an inhibitor
CC protein of nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B
CC degradation via the ubiquitin pathway is useful for identifying
CC modulators of this process for use in treating diseases associated with
CC activation of NF-kappa-B. In vitro analysis suggests that deletion of
CC the F-box results in a protein that functions as a dominant negative
CC molecule in vivo. Transient over-expression of delta-beta-Trcp (a
CC deletion mutant) inhibited the degradation of endogenous I-kappa-B-alpha
CC in stimulated Jurkat cells, resulting in accumulation of phosphorylated
CC I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B
CC activity. E3 and beta-Trcp can be used to modulate NF-kappa-B to treat
CC inflammatory diseases, autoimmune diseases, cancer and viral infections.
XX
SO Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

Query Match 100.0%; Score 2151; DB 21; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGGTTGGCTGGCGCCGCGCACCACCAAGGGGGCGCCGCGGAGAGCCAGTGGCC 60
    |||||||
Db 1 TGGGTTGGCTGGCGCCGCGCACCACCAAGGGGGCGCCGCGGAGAGCCAGTGGCC 60
QY 61 TGGGCGATTATGAGCCCGCGGAGGGGGTGTCTGCAAGAGAGGCACTCAATTATGAT 120
    |||||||
Db 61 TGGGCGATTATGAGCCCGCGGAGGGGGTGTCTGCAAGAGAGGCACTCAATTATGAT 120
QY 121 TCCTCAGAGAGAGAACTATTAATGCGGAACCCCTCAGCAAGTAATACCAAGAG 180
    |||||||
Db 121 TCCTCAGAGAGAGAACTATTAATGCGGAACCCCTCAGCAAGTAATACCAAGAG 180
QY 181 AATTCATCTAGACAGATCAACACAGCTGTGCCAGACTGTCTTAACCAAGAAACAGTA 240
    |||||||
Db 181 AATTCATCTAGACAGATCAACACAGCTGTGCCAGACTGTCTTAACCAAGAAACAGTA 240
QY 241 TGTTCAGACAGCACTGTATGAAGACTGAGAAATTTGTGGCCAAAACAAACTTGCAT 300
    |||||||
Db 241 TGTTCAGACAGCACTGTATGAAGACTGAGAAATTTGTGGCCAAAACAAACTTGCAT 300
QY 301 GGCACCTTCCAGTATGATTTGGCCCAAGCAAGGAAACTCTCAGCAAGTATGAAAAGAA 360
    |||||||
Db 301 GGCACCTTCCAGTATGATTTGGCCCAAGCAAGGAAACTCTCAGCAAGTATGAAAAGAA 360
QY 361 AAGGAAGTGTGTCAAAATCTTGAAGAGTGTGAGAGTCAGATCAAGTGAATTTGTG 420
    |||||||
Db 361 AAGGAAGTGTGTCAAAATCTTGAAGAGTGTGAGAGTCAGATCAAGTGAATTTGTG 420
QY 421 GAACATCTTATCCCAATGCTCTATACCAATGAGGACATAAATCTGATCTTAA 480
    |||||||
Db 421 GAACATCTTATCCCAATGCTCTATACCAATGAGGACATAAATCTGATCTTAA 480
QY 481 CCTATGTCAGAGAGATTTCAATACCTCTGCGCAGCTCGGGGATGGATATATCGCT 540
    |||||||
Db 481 CCTATGTCAGAGAGATTTCAATACCTCTGCGCAGCTCGGGGATGGATATATCGCT 540
QY 541 GAGAACCTTGTGCATACCGGATGCAATCTATGTCGTCTGAATGTCGCAAG 600
    |||||||

```

```

Db 541 GAGACATTTCTGTACATACCTGATGCGCAAAATCATATGTGCTGACCTTGTCGAAG 600
QY 601 GAATGATACGAGTGCACCTGTGATGGCATGTGTGGAGAAGCTTATTCAGAGATATGTC 660
Db 601 GAATGATACGAGTGCACCTGTGATGGCATGTGTGGAGAAGCTTATTCAGAGATATGTC 660
QY 661 AGCAGACATTTCTGTGTGGAGAGGCTTGCGACAAAGAGAGATGGGAGACATATTTATTC 720
Db 661 AGCAGACATTTCTGTGTGGAGAGGCTTGCGACAAAGAGAGATGGGAGACATATTTATTC 720
QY 721 AAAAAAATCTCTGAGCGGGAATGCTCTCCCACTCTTTTATGTAGACACTTTATCTCT 780
Db 721 AAAAAAATCTCTGAGCGGGAATGCTCTCCCACTCTTTTATGTAGACACTTTATCTCT 780
QY 781 AAAAAATATACAGACATTTGAGACATATGATCTAATTTGGAGATGTGGAGACATAGTTTA 840
Db 781 AAAAAATATACAGACATTTGAGACATATGATCTAATTTGGAGATGTGGAGACATAGTTTA 840
QY 841 CAGAGACATTTCTGTGCGAGAGTGAACAGCAAGAGCTTACTGTTTACAGTATGATGAT 900
Db 841 CAGAGACATTTCTGTGCGAGAGTGAACAGCAAGAGCTTACTGTTTACAGTATGATGAT 900
QY 901 CAGAAAAATGTAAGCGGCTTGCGAGACAAACACATCAAGATCGGGATAAAAACACATTG 960
Db 901 CAGAAAAATGTAAGCGGCTTGCGAGACAAACACATCAAGATCGGGATAAAAACACATTG 960
QY 961 GAATGCAAGCAATTTCTCAGAGGCTACAGATGAGTTCAGTCTGTCCTGTCAGATGATGAT 1020
Db 961 GAATGCAAGCAATTTCTCAGAGGCTACAGATGAGTTCAGTCTGTCCTGTCAGATGATGAT 1020
QY 1021 AGAGTGTCTAATCAGAGATCATGGATTCAGAGGCTCAGAGATGGAGATGTAATACAGT 1080
Db 1021 AGAGTGTCTAATCAGAGATCATGGATTCAGAGGCTCAGAGATGGAGATGTAATACAGT 1080
QY 1081 GAAATGCTAATCAACAGTGTGATTCACATTTGAGCAAGTTCAGTCTGCTGCTTCAATAT 1140
Db 1081 GAAATGCTAATCAACAGTGTGATTCACATTTGAGCAAGTTCAGTCTGCTGCTTCAATAT 1140
QY 1141 GGCATGATGCTGACCTGCTCCAAAGATGCTCATGCTGATGGATATGGGCTCCCA 1200
Db 1141 GGCATGATGCTGACCTGCTCCAAAGATGCTCATGCTGATGGATATGGGCTCCCA 1200
QY 1201 ACTGACATTTACCTCCGAGAGGCTGCTGTCGACACCGAGCTGCTCATGTTGATGAC 1260
Db 1201 ACTGACATTTACCTCCGAGAGGCTGCTGTCGACACCGAGCTGCTCATGTTGATGAC 1260
QY 1261 TTTGATGACACAGTACATTTGTTTCTGATCTGGGATAGACATTAAGGTATGAGACACA 1320
Db 1261 TTTGATGACACAGTACATTTGTTTCTGATCTGGGATAGACATTAAGGTATGAGACACA 1320
QY 1321 AGTACTTGTGAATTTGTAGAGACCTTAATGACACAAAGAGGCTTCCCTGTTTGACG 1380
Db 1321 AGTACTTGTGAATTTGTAGAGACCTTAATGACACAAAGAGGCTTCCCTGTTTGACG 1380
QY 1381 TACAGGACAGGCTGCTGATGAGTGTGCTCATGACACATCATGATGATGAGACATA 1440
Db 1381 TACAGGACAGGCTGCTGATGAGTGTGCTCATGACACATCATGATGATGAGACATA 1440
QY 1441 GAATGATGATGATTTAGAGATGTTAGAGAGGCTATGAGAGGATTTGGTGTGATGCA 1500
Db 1441 GAATGATGATGATTTAGAGATGTTAGAGAGGCTATGAGAGGATTTGGTGTGATGCA 1500
QY 1501 TTTGATTAACAGAGATGATGATGAGGCTTATGATGAGAAATTAATTAATGATGATCTT 1560
Db 1501 TTTGATTAACAGAGATGATGATGAGGCTTATGATGAGAAATTAATTAATGATGATCTT 1560
QY 1561 GTGGCTGCTTTTGGACCCCGCTGCTCTGCAAGGACACTGTGCTACAGACCTTGTGGAG 1620
Db 1561 GTGGCTGCTTTTGGACCCCGCTGCTCTGCAAGGACACTGTGCTACAGACCTTGTGGAG 1620
QY 1621 CATTCGGAAGAGATTTTGGACTAGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1621 CATTCGGAAGAGATTTTGGACTAGATGATGATGATGATGATGATGATGATGATGATGAT 1680

```

```

QY 1681 GATGACACATTCATTCATTCGAGGACTTCCTAATATATCATCAGCTGCCAAGCTGAACCCCC 1740
Db 1681 GATGACACATTCATTCATTCGAGGACTTCCTAATATATCATCAGCTGCCAAGCTGAACCCCC 1740
QY 1741 CGTTCCTCTTCGCAACATACACCTACATTCCTCAGATTAATTAATTAATTAATTAATTAAT 1800
Db 1741 CGTTCCTCTTCGCAACATACACCTACATTCCTCAGATTAATTAATTAATTAATTAATTAAT 1800
QY 1801 ACTTCCCGAGAGACCATTAAGTTGGGCTATTTAAGTATCTGCAATGATGATGATGATGATG 1860
Db 1801 ACTTCCCGAGAGACCATTAAGTTGGGCTATTTAAGTATCTGCAATGATGATGATGATGATG 1860
QY 1861 AACACAGTAACTCAAACTACTGCCAGTTTCCCTGAGATGCCAGAGAGAGGCTT 1920
Db 1861 AACACAGTAACTCAAACTACTGCCAGTTTCCCTGAGATGCCAGAGAGAGGCTT 1920
QY 1921 TGAGACTCTCTGTTGGGACACAGTTGCTGCAAGTGGGCGCCAGAGAGGCTTACTAGCACA 1980
Db 1921 TGAGACTCTCTGTTGGGACACAGTTGCTGCAAGTGGGCGCCAGAGAGGCTTACTAGCACA 1980
QY 1981 ACTGACTCTCTGAGTGTGCTATCAGAAAGATGCTCTGATCAATTTGAAATGATGGAAC 2040
Db 1981 ACTGACTCTCTGAGTGTGCTATCAGAAAGATGCTCTGATCAATTTGAAATGATGGAAC 2040
QY 2041 TTTTAACTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
Db 2041 TTTTAACTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
QY 2101 CAGACAAAGGAGCTTAAATATATATATATATATATATATATATATATATATATATATAT 2151
Db 2101 CAGACAAAGGAGCTTAAATATATATATATATATATATATATATATATATATATATATAT 2151

RESULT 3
AAZ93350
ID AAZ93350 standard; cdna; 2151 bp.
XX
AC AAZ93350;
XX
DT 16-AUG-2000 (first entry)
XX
DE Sequence encoding F-box protein FBP-1.
XX
KW F-box protein; FBP; diagnosis; treatment; screening; agonist;
KW antagonist; proliferative disorder; differential disorder;
KW breast cancer; prostate cancer; ovarian cancer; cancer;
KW small cell lung carcinoma; immune disorder; cardiovascular disorder;
KW inflammatory disorder; human; ss.
XX
OS Homo sapiens.
XX
FH Key 70..1779 Location/Qualifiers
FT CDS /*tag= a
FT FT /product= FBP-1
FT FT
PN MO200012679-A1.
XX
PD 09-MAR-2000.
XX
PE 27-AUG-1999; 99MO-US19560.
XX
PR 28-AUG-1998; 98US-0098355.
PR 03-FEB-1999; 99US-0118568.
PR 15-MAR-1999; 99US-0124449.
XX
PA (UNY ) UNIV NEW YORK STATE.
XX
PI Chlaure DS, Pagano M, Latres E;
XX
DR WPI: 2000-256635/22.
DR P-PSDB; AAB83041.

```

xx Novel nucleic acid for screening compounds useful for treating  
 PT proliferative and differentiative disorders such as cancer and immune  
 PT disorders comprises sequences encoding ubiquitin ligases  
 xx  
 PS Disclosure; Figure 3b, 3c, 3d; 245pp; English.  
 CC  
 CC Nucleic acids encoding substrate-targeting subunits of ubiquitin  
 CC ligases with F-box motifs (F-box proteins) are useful for diagnosis  
 CC of proliferative and differentiated related disorders by measuring  
 CC FBP gene expression. Cells expressing such proteins or  
 CC their fragments are useful for screening compounds. The compounds  
 CC are agonists or antagonists, which are useful for treating a  
 CC proliferative or differentiative disorder in a mammal such as  
 CC breast, ovarian and prostate cancer and small cell lung carcinoma  
 CC and also major opportunistic infections, immune disorders,  
 CC cardiovascular diseases and inflammatory disorders. FBP protein,  
 CC analogs, derivatives and their subsequences, anti-FBP antibodies  
 CC are also useful in diagnosis of the disorders.  
 CC  
 SQ Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

Query Match 100.0%; Score 2151; DB 21; Length 2151;  
 Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;  
 Matches 2151; Conservative 0; Indels 0;

QY 1 TCGTTGGCTGGCGCTGGACCAAGGCGCGCGCGGAGAGCGACCACTGGCC 60  
 DB 1 TCGTTGGCTGGCGCTGGACCAAGGCGCGCGCGGAGAGCGACCACTGGCC 60  
 QY 61 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 120  
 DB 61 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 120  
 QY 121 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 180  
 DB 121 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 180  
 QY 181 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 240  
 DB 181 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 240  
 QY 241 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 300  
 DB 241 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 300  
 QY 301 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 360  
 DB 301 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 360  
 QY 361 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 420  
 DB 361 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 420  
 QY 421 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 480  
 DB 421 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 480  
 QY 481 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 540  
 DB 481 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 540  
 QY 541 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 600  
 DB 541 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 600  
 QY 601 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 660  
 DB 601 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 660  
 QY 661 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 720  
 DB 661 TCGCGGATTGAGACCGCGCGCGCGCGCGCGCGGAGAGCGACCACTGGCC 720

QY 721 AAAAACAACCTCTGACGGGAAATGCTCTCCCACTCTTTTATAGACACTTTATCC 780  
 DB 721 AAAAACAACCTCTGACGGGAAATGCTCTCCCACTCTTTTATAGACACTTTATCC 780  
 QY 781 AAAATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 840  
 DB 781 AAAATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 840  
 QY 841 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 900  
 DB 841 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 900  
 QY 901 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 960  
 DB 901 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 960  
 QY 961 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1020  
 DB 961 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1020  
 QY 1021 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1080  
 DB 1021 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1080  
 QY 1081 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1140  
 DB 1081 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1140  
 QY 1141 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1200  
 DB 1141 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1200  
 QY 1201 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1260  
 DB 1201 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1260  
 QY 1261 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1320  
 DB 1261 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1320  
 QY 1321 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1380  
 DB 1321 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1380  
 QY 1381 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1440  
 DB 1381 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1440  
 QY 1441 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1500  
 DB 1441 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1500  
 QY 1501 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1560  
 DB 1501 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1560  
 QY 1561 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1620  
 DB 1561 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1620  
 QY 1621 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1680  
 DB 1621 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1680  
 QY 1681 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1740  
 DB 1681 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1740  
 QY 1741 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1800  
 DB 1741 CAGAGATTTACAGACATTTGACATAGATTAATTTGAGATGAGAGACATATCTTA 1800







ID AAH90079 standard; cDNA; 3220 BP.  
XX  
AC AAH90079;  
XX  
DT 01-OCT-2001 (first entry)  
XX  
DE Human bone marrow cDNA, SEQ ID NO: 323.  
XX  
KW Human; bone marrow; antineoplastic; cytostatic; neuroprotective;  
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;  
KW immunosuppressive; gene therapy; cytokine cell proliferation;  
KW cell differentiation modulator; immune disorder; infection; cancer;  
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.  
XX  
OS Homo sapiens.  
XX  
PN W0200153453-A2.  
XX  
PD 26-JUL-2001.  
XX  
PF 23-DEC-2000; 2000WO-US34960.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
XX  
PR 25-APR-2000; 2000US-0552317.  
XX  
PR 09-JUL-2000; 2000US-0598042.  
XX  
PR 19-JUL-2000; 2000US-0620312.  
XX  
PR 03-AUG-2000; 2000US-0634350.  
XX  
PR 14-SEP-2000; 2000US-0662191.  
XX  
PR 19-OCT-2000; 2000US-0693036.  
XX  
PR 30-NOV-2000; 2000US-0250583.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;  
PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Drmanac RT;  
XX  
DR WPI: 2001-488707/53.  
XX  
DR P-PSDB: AAM00960.  
XX  
PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful  
PT for treating e.g. cancer and immune deficiency disorders -  
XX  
PS Claim 1; Page 428; 648pp; English.  
XX  
CC The present sequence is one of 251 novel human polynucleotides  
CC expressed in the bone marrow. The polynucleotide and the  
CC polypeptide encoded by it are useful in the treatment of various  
CC immune deficiencies and disorders. The deficiencies and disorders may  
CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal  
CC infection, or may result from an autoimmune disorder, a coagulation  
CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,  
CC suppression of an inflammatory response or treatment of a nervous  
CC system disorder such as Alzheimer's disease. Detection of the presence  
CC or increased expression of the polynucleotide or the protein it  
CC encodes is useful for the diagnosis and/or prognosis of one  
CC or more types of cancer. The polynucleotide and polypeptide can be  
CC used as nutritional sources or supplements and in the screening of  
CC chemical compounds as potential drugs.  
XX  
SQ Sequence 3220 BP; 878 A; 725 C; 770 G; 847 T; 0 other;

Query Match 98.6%; Score 2120.8; DB 22; Length 3220;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 2136; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1 TGGCGTGGCTGGCGGACCAAAAGGGGGGGCCCGCGGAGAGCGGACCGCTGGCC 60  
DB TGGCTGGCTGGCGGCTGGCGGACCAAAAGGGGGGGCCCGCGGAGAGCGGACCGCTGGCC 118  
QY 61 TCGGCGATTATGAGACCGCGGAGCGGCTGCTGCAAGAGAGCACTCAAGTTATGAT 120  
DB TCGGCGATTATGAGACCGCGGAGCGGCTGCTGCAAGAGAGCACTCAAGTTATGAT 178

QY 121 TCCTCAGAGAGAGAGAGACTGTAAATATGGGAAACCCCTAGAGAGATTAATACCAGAGAG 180  
DB TCCTCAGAGAGAGAGAGAGACTGTAAATATGGGAAACCCCTAGAGAGATTAATACCAGAGAG 238  
QY 181 AATTTCATTAGACAGACATTAACACAGCTGTCCAGACTCTGCTTAACCAAGAACAGTA 240  
DB AATTTCATTAGACAGACATTAACACAGCTGTCCAGACTCTGCTTAACCAAGAACAGTA 298  
QY 241 TGTTCAGCAAGCACTCTATGAGAGAGTGAATGTGTGGCCAAACAAACAACTGGCAAT 300  
DB TGTTCAGCAAGCACTCTATGAGAGAGTGAATGTGTGGCCAAACAAACAACTGGCAAT 358  
QY 301 GGCACCTTCAGATATGATTTGTGCCCAAGCAACGAAACTCTCAGACAGCTATGAAAAAGAA 360  
DB GGCACCTTCAGATATGATTTGTGCCCAAGCAACGAAACTCTCAGACAGCTATGAAAAAGAA 418  
QY 361 AAGGAATGCTGTGCAAAATCTTTGAGAGAGTGTGACAGATCAGATCAAGTGAATTTG 420  
DB AAGGAATGCTGTGCAAAATCTTTGAGAGAGTGTGACAGATCAGATCAAGTGAATTTG 478  
QY 421 GAACATCTTATATCCCAAAATGTGTCAATTAACACATGAGGACATTAACCTGATCTTAA 480  
DB GAACATCTTATATCCCAAAATGTGTCAATTAACACATGAGGACATTAACCTGATCTTAA 538  
QY 481 CCTATGTTGCAAGAGATTTCAATTAAGTCTGCTGCAAGCTGGGAGATTTGATATGCT 540  
DB CCTATGTTGCAAGAGATTTCAATTAAGTCTGCTGCAAGCTGGGAGATTTGATATGCT 598  
QY 541 GAGAACATCTGTGCAATTAAGTGTGCAAAATCTATGAGAGAGTGTGAGTGTGCAAG 600  
DB GAGAACATCTGTGCAATTAAGTGTGCAAAATCTATGAGAGAGTGTGAGTGTGCAAG 658  
QY 601 GAATGTACAGAGTGAAGCTGTGATGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 660  
DB GAATGTACAGAGTGAAGCTGTGATGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 718  
QY 661 AGGACAGATTTCTGTGAGAGAGGCTGTGCAAGAGAGAGAGTGTGAGAGAGTGTGAG 720  
DB AGGACAGATTTCTGTGAGAGAGGCTGTGCAAGAGAGAGAGTGTGAGAGAGTGTGAG 778  
QY 721 AAAAAGCAACCTCTGTGAGAGAGGAGTGTGCAAGAGAGAGTGTGAGAGAGTGTGAG 780  
DB AAAAAGCAACCTCTGTGAGAGAGGAGTGTGCAAGAGAGAGTGTGAGAGAGTGTGAG 838  
QY 781 AAAATTATACAGACATTTGAGACATTAAGATTAATGAGATGTGAGAGATTTA 840  
DB AAAATTATACAGACATTTGAGACATTAAGATTAATGAGATGTGAGAGATTTA 898  
QY 841 CAGAGATTTCTGTGAGAGAGGCTGTGCAAGAGAGAGAGTGTGAGAGAGTGTGAG 900  
DB CAGAGATTTCTGTGAGAGAGGCTGTGCAAGAGAGAGAGTGTGAGAGAGTGTGAG 958  
QY 901 CAGAGATTTCTGTGAGAGAGGCTGTGCAAGAGAGAGAGTGTGAGAGAGTGTGAG 960  
DB CAGAGATTTCTGTGAGAGAGGCTGTGCAAGAGAGAGAGTGTGAGAGAGTGTGAG 1018  
QY 961 GAATGCAAGGAAATTTCTACAGAGGCAATACAGTTCACTGCTGTCCAGTATGATGAG 1020  
DB GAATGCAAGGCAATTTCTACAGAGGCAATACAGTTCACTGCTGTCCAGTATGATGAG 1078  
QY 1021 AGAGTATCTAATACAGATATGAGATTTCAAGGCTGTGAGAGTGTGAGATTTA 1080  
DB AGAGTATCTAATACAGATATGAGATTTCAAGGCTGTGAGAGTGTGAGATTTA 1138  
QY 1081 GAATGCTAATACAGATATGAGATTTCAAGGCTGTGAGAGTGTGAGATTTA 1140  
DB GAATGCTAATACAGATATGAGATTTCAAGGCTGTGAGAGTGTGAGATTTA 1198  
QY 1141 GGCATGATGTGAGCTGTCTCAAAAGATCGTTCCATTGCTGTATGGATATGGCTTCCCA 1200  
DB GGCATGATGTGAGCTGTCTCAAAAGATCGTTCCATTGCTGTATGGATATGGCTTCCCA 1258

QY 1201 ACTGACATTACCTCCGGAGGGTGTGTCGAGACAGGAGCTGTCAATGTTGAGAC 1260  
 |||  
 DB 1259 ACTGACATTACCTCCGGAGGGTGTGTCGAGACAGGAGCTGTCAATGTTGAGAC 1318  
 QY 1261 TTTGATGACAAAGTACATTTGTTCTGCATCTGGGGATAGAATTAAGCTATGAGACAA 1320  
 |||  
 DB 1319 TTTGATGACAAAGTACATTTGTTCTGCATCTGGGGATAGAATTAAGCTATGAGACAA 1378  
 QY 1321 AGTACTTGTGATTTGTTAAGGACCTTAATGACACAAAGGAGCATGCTGTTTGACG 1380  
 |||  
 DB 1379 AGTACTTGTGATTTGTTAAGGACCTTAATGACACAAAGGAGCATGCTGTTTGACG 1438  
 QY 1381 TACAGGACAGGCTGTGATGAGTGCATCTGACAAACATATCATGATTATGAGACATA 1440  
 |||  
 DB 1439 TACAGGACAGGCTGTGATGAGTGCATCTGACAAACATATCATGATTATGAGACATA 1498  
 QY 1441 GAAATGTGTGATTTTACAGTGTAGAAAGCCATGAGCAATGCTGCTTATTCGA 1500  
 |||  
 DB 1499 GAAATGTGTGATTTTACAGTGTAGAAAGCCATGAGCAATGCTGCTTATTCGA 1558  
 QY 1501 TTTGATTAACAAGAGATAGTACAGTGGGCTATGATGAGAAAAATTAAAGTGTGACATCT 1560  
 |||  
 DB 1559 TTTGATTAACAAGAGATAGTACAGTGGGCTATGATGAGAAAAATTAAAGTGTGACATCT 1618  
 QY 1561 GTGGCTGCTTTGAGACCCGCTGCTCGAGAGGACACTGTGCTACGAGACCCCTTGAGAG 1620  
 |||  
 DB 1619 GTGGCTGCTTTGAGACCCGCTGCTCGAGAGGACACTGTGCTACGAGACCCCTTGAGAG 1678  
 QY 1621 CATTTCCGGAAGAGTTTTCGACATACAGTTTGATGAATTCAGATTCGAGTGTACAT 1680  
 |||  
 DB 1679 CATTTCCGGAAGAGTTTTCGACATACAGTTTGATGAATTCGAGTGTGTACAT 1738  
 QY 1681 GATGACACATCTCTATGTCGGGACTCTAAATGATTCAGCGCCCAACCTTAACCCGCC 1740  
 |||  
 DB 1739 GATGACACATCTCTATGTCGGGACTCTAAATGATTCAGCGCCCAACCTTAACCCGCC 1798  
 QY 1741 CGTTCCTCTTGGAAACATACACCTCATCTCCAGATAAATAACATACATGACCTCAT 1800  
 |||  
 DB 1799 CGTTCCTCTTGGAAACATACACCTCATCTCCAGATAAATAACATACATGACCTCAT 1858  
 QY 1801 ACTTGGCCAGGAGCCCAATTAAG--TTGGGGATTTTAACGTAATTCGCCAATACCAAGATGAG 1859  
 |||  
 DB 1859 ACTTGGCCAGGAGCCCAATTAAGTTTGGGATTTTAACGTAATTCGCCAATACCAAGATGAG 1918  
 QY 1860 CAACAACAGTAACAACTACTGCCAGTTTCCCTGGAGTAGCCGAGAGCAAGGCT 1919  
 |||  
 DB 1919 CAACAACAGTAACAACTACTGCCAGTTTCCCTGGAGTAGCCGAGAGCAAGGCT 1978  
 QY 1920 TTGAGACTCTGTGGGACACAGTGTGTGCAAGTGGCCGAGAGCGTCTACTACGAC 1979  
 |||  
 DB 1979 TTGAGACTCTGTGGGACACAGTGTGTGCAAGTGGCCGAGAGCGTCTACTACGAC 2038  
 QY 1980 AACTGACTCTTCAAGTGTGTATCAGAAAGATGTCTTATCAATTAATTAAGATTTGAA 2039  
 |||  
 DB 2039 AACTGACTCTTCAAGTGTGTATCAGAAAGATGTCTTATCAATTAATTAAGATTTGAA 2098  
 QY 2040 CTTTAAACCTCCCTCCTCCTCCTCCTTACCTTGACCTAGTATTTTCCATTTGTT 2099  
 |||  
 DB 2099 CTTTAAACCTCCCTCCTCCTCCTCCTTACCTTGACCTAGTATTTTCCATTTGTT 2158  
 QY 2100 CCAGCAAAAGTGACTTATTAATTTATTTAGTGTGTTGCAAAA 2143  
 |||  
 DB 2159 CCAGCAAAAGTGACTTATTAATTTATTTAGTGTGTTGCAAAA 2202

DE Human cell signalling protein-12 encoding cDNA.  
 XX  
 KW Cell signalling protein-12; CSIGP-12; cell proliferation;  
 KW inflammatory disorder; cirrhosis; cancer; hepatitis; AIDS;  
 KW arteriosclerosis; Addison's disease; multiple sclerosis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 70..1779  
 FT /tag= a  
 FT /product= "Cell Signalling protein-12"  
 XX  
 PN W0958558-A2.  
 XX  
 PD 18-NOV-1999.  
 XX  
 PF 13-MAY-1999; 99WO-US10567.  
 XX  
 PR 13-MAY-1998; 98US-0085343.  
 PR 26-AUG-1998; 98US-0098010.  
 XX  
 PA (INCYTE) INCYTE PHARM INC.  
 XX  
 PI Bandman O, Hillman JL, Lal P, Yue H, Tang YF, Patterson C;  
 PI Baughn MR, Yang J;  
 XX  
 DR WPI: 2000-086432/07.  
 DR P-PSDB; AAY44249.  
 XX  
 PT Human cell signalling proteins useful for, e.g. diagnosing cell  
 PT proliferative and inflammatory disorders -  
 PS  
 PS Claim 9; Page 87-88; 90pp; English.  
 XX  
 CC The present sequence is a cDNA obtained from Incyte clone 3239149 of  
 CC COA00T01 library. It encodes cell signalling protein-12 (CSIGP-12). It  
 CC is expressed in musculo-skeletal, gastrointestinal and nervous  
 CC tissues. Fragments of CSIGP encoding nucleic acid can be used as  
 CC hybridisation probe for detecting CSIGP related sequences or allelic  
 CC variants. Recombinant CSIGP can be produced in host cells by transforming  
 CC them with genetically engineered vectors. Agonists or antagonists can be  
 CC used in the treatment of cell proliferative and inflammatory disorders  
 CC associated with decreased or increased CSIGP expression. CSIGP is used in  
 CC the diagnosis, prevention and treatment of cell proliferative disorders  
 CC like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory  
 CC disorders like AIDS, Addison's disease, multiple sclerosis, etc.  
 XX  
 SO Sequence 2419 BP; 671 A; 531 C; 625 G; 590 T; 2 other;  
 Query Match 98.3%; Score 2114.8; DB 21; Length 2419;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 2119; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 11 GGGGCTTGACCAAAAGGGGCGCCGCGGAGAGGAGCCAGTGGCTGCGGATTA 70  
 |||  
 DB 11 GGGGCTTGACCAAAAGGGGCGCCGCGGAGAGGAGCCAGTGGCTGCGGATTA 70  
 QY 71 TGGACCCGGCGGAGGCGGTGTGCAAGAAAGGACATCAAGTTTATTAATTCCTCAGAGA 130  
 |||  
 DB 71 TGGACCCGGCGGAGGCGGTGTGCAAGAAAGGACATCAAGTTTATTAATTCCTCAGAGA 130  
 QY 131 GAGAAAGTGTAAATTAATGAGAAACCCCTAGAAATATACAGAGAAATTCACCTTA 190  
 |||  
 DB 131 GAGAAAGTGTAAATTAATGAGAAACCCCTAGAAATATATACAGAGAAATTCACCTTA 190  
 QY 191 GACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTATGTTAGCAA 250  
 |||  
 DB 191 GACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTATGTTAGCAA 250  
 QY 251 GCACGTGATGAAGACATGAGAAATTTGTGGCCAAAACAAACTTGGCAATGGACATTCCA 310  
 |||  
 DB 251 GCACGTGATGAAGACATGAGAAATTTGTGGCCAAAACAAACTTGGCAATGGACATTCCA 310

RESULT 7  
 AA229233  
 ID AA229233 standard; cDNA; 2419 BP.  
 XX  
 AC AA229233;  
 XX  
 DT 28-FEB-2000 (first entry)  
 XX



PD 26-JUL-2001.  
 XX  
 PF 23-DEC-2000; 2000MO-US34960.  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 30-NOV-2000; 2000US-0250583.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;  
 PI Ren F, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Drmanac RT;  
 XX  
 DR WPI: 2001-488707/53.  
 DR P-PSDB; AAM00847.  
 XX  
 PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful  
 PT for treating e.g. cancer and immune deficiency disorders -  
 XX  
 PS Claim 1: Page 274-275; 648pp; English.  
 XX  
 CC The present sequence is one of 251 novel human polynucleotides  
 CC expressed in the bone marrow. The polynucleotide and the  
 CC polypeptide encoded by it are useful in the treatment of various  
 CC immune deficiencies and disorders. The deficiencies and disorders may  
 CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal  
 CC infection, or may result from an autoimmune disorder, a coagulation  
 CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,  
 CC suppression of an inflammatory response or treatment of a nervous  
 CC system disorder such as Alzheimer's disease. Detection of the presence  
 CC or increased expression of the polynucleotide or the protein it  
 CC encodes is useful for the diagnosis and/or prognosis of one  
 CC or more types of cancer. The polynucleotide and polypeptide can be  
 CC used as nutritional sources or supplements and in the screening of  
 CC chemical compounds as potential drugs.  
 XX  
 SQ Sequence 3622 BP; 961 A; 839 C; 874 G; 948 T; 0 other;  
 Query Match 93.5%; Score 2011.4; DB 22; Length 3622;  
 Best local similarity 99.7%; Pred. No. 0;  
 Matches 2026; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Db 621 TATCTTAACCTATGTTGACAGAGATTTTATATACCTCTGCGACGCTGGGGATTGGAT 680  
 |||||||  
 QY 532 CATATCGCTGAGAACATTCGTCTCATACCTGGATGCGCAATCACTATGCTGCTGAACCT 591  
 |||||||  
 Db 681 CATATCGCTGAGAACATTCGTCTCATACCTGGATGCGCAATCACTATGCTGCTGAACCT 740  
 |||||||  
 QY 592 GTGTGCAAGGATGATGACCGAGTACCTCTGATGGCATGCTGTGGGAAGAGCTTATCGAG 651  
 |||||||  
 Db 741 GTGTGCAAGGATGATGACCGAGTACCTCTGATGGCATGCTGTGGGAAGAGCTTATCGAG 800  
 |||||||  
 QY 652 AGAATGTCAGACAGATTCCTGTTGAGAGGCGCTGCGAAGAGAGATGGGAGCAG 711  
 |||||||  
 Db 801 AGAATGTCAGACAGATTCCTGTTGAGAGGCGCTGCGAAGAGAGATGGGAGCAG 860  
 |||||||  
 QY 712 TATTTATTCACAAACAAACCTCTGACGGGAGTCTCTCCCACTCTTTTATAGACA 771  
 |||||||  
 Db 861 TATTTATTCACAAACAAACCTCTGACGGGAGTCTCTCCCACTCTTTTATAGACA 920  
 |||||||  
 QY 772 CTTTATCCTAAATTTATACAAAGACATGAGACATTAATCTAATGAGATGAGAGAGA 831  
 |||||||  
 Db 921 CTTTATCCTAAATTTATACAAAGACATGAGACATTAATCTAATGAGATGAGAGAGA 980  
 |||||||  
 QY 832 CATAGTTTACAGAGATTCACCTGCGAAGTGAACAAAGAGATTTACTGTTTACAG 891  
 |||||||  
 Db 981 CATAGTTTACAGAGATTCACCTGCGAAGTGAACAAAGAGATTTACTGTTTACAG 1040  
 |||||||  
 QY 892 TATGATGATCAAAAATAGTAAAGCGGCTTCGAGACACACATCAAGATCTGGATAAA 951  
 |||||||  
 Db 1041 TATGATGATCAAAAATAGTAAAGCGGCTTCGAGACACACATCAAGATCTGGATAAA 1100  
 |||||||  
 QY 952 AACACATGGAATGCAAGCGAATTCACAGAGCCATCAGATTCAGTCTGCTCAG 1011  
 |||||||  
 Db 1101 AACACATGGAATGCAAGCGAATTCACAGAGCCATCAGATTCAGTCTGCTCAG 1160  
 |||||||  
 QY 1012 TATGATGAGAGATGATCATTAACAGATCATGAGATTCACAGTGTGGATGTA 1071  
 |||||||  
 Db 1161 TATGATGAGAGATGATCATTAACAGATCATGAGATTCACAGTGTGGATGTA 1220  
 |||||||  
 QY 1072 AATACAGGTGAATGCTTAAACAGCGTGAATCACCATTGTGAAGAGATTCGACATTCG 1131  
 |||||||  
 Db 1221 AATACAGGTGAATGCTTAAACAGCGTGAATCACCATTGTGAAGAGATTCGACATTCG 1280  
 |||||||  
 QY 1132 TTCAATAATGCGATGATGATGACCTGCTCCAAAGATGCTTCATGCTGATGGATATG 1191  
 |||||||  
 Db 1281 TTCAATAATGCGATGATGATGACCTGCTCCAAAGATGCTTCATGCTGATGGATATG 1340  
 |||||||  
 QY 1192 GCGTCCCACTGACATTAACCTCCGAGAGGTGCTGTGCGACACGAGCTGCTGCAAT 1251  
 |||||||  
 Db 1341 GCGTCCCACTGACATTAACCTCCGAGAGGTGCTGTGCGACACGAGCTGCTGCAAT 1400  
 |||||||  
 QY 1252 GTTTAGACTTTGATGACAACTACATTTGTTCTGCATCTGGGGATAGAACTATAAAGTA 1311  
 |||||||  
 Db 1401 GTTTAGACTTTGATGACAACTACATTTGTTCTGCATCTGGGGATAGAACTATAAAGTA 1460  
 |||||||  
 QY 1312 TGGAAACACAGTACTGTTGAATTTGTAAAGACCTTAAATGAGACAAACAGAGCATTC 1371  
 |||||||  
 Db 1461 TGGAAACACAGTACTGTTGAATTTGTAAAGACCTTAAATGAGACAAACAGAGCATTC 1520  
 |||||||  
 QY 1372 TGTGTCAGTACAGGACAGGCTGGTGTGATGAGTGCATGTGACACACATCATGATTA 1431  
 |||||||  
 Db 1521 TGTGTCAGTACAGGACAGGCTGGTGTGATGAGTGCATGTGACACACATCATGATTA 1480  
 |||||||  
 QY 1432 TGGGACATAGATGATGCTGATGCTTTTACAGTGTAGAAAGCCATGAGAAATGGTCCGT 1491  
 |||||||  
 Db 1581 TGGGACATAGATGATGCTGATGCTTTTACAGTGTAGAAAGCCATGAGAAATGGTCCGT 1640  
 |||||||  
 QY 1492 TGTATTCATTTGATTAACAAAGATGATGCTGAGGCGCTATGATGAAATAATTAAGTG 1551  
 |||||||  
 Db 1641 TGTATTCATTTGATTAACAAAGATGATGCTGAGGCGCTATGATGAAATAATTAAGTG 1700  
 |||||||  
 QY 1552 TGGGATCTTGTGGGCTGCTTGGAGCCCGCTGCTGACGAGGACACGCTGCTACGAGCC 1611  
 |||||||

Db 1701 TGGGATCTTGGGCTGGCTTTGACACCCCGTCTGCTGACAGGACACTCTGTCTACGGACC 1760  
 QY 1612 CTGTGTGAGCATTCGCGAGAGATTTTTCGACTACAGTTTGATGTAATTCAGATTGTCAAT 1671  
 Db 1761 CTGTGTGAGCATTCGCGAGAGATTTTTCGACTACAGTTTGATGTAATTCAGATTGTCAAT 1820  
 QY 1672 AGTTACATGATGACAAATCCCATCTGGAGTTTCTTAATATCCAGCTGCCAAGCT 1731  
 Db 1821 AGTTACATGATGACAAATCCCATCTGGAGTTTCTTAATATCCAGCTGCCAAGCT 1880  
 QY 1732 GAACCCCCCGGTTCCCTTCTCGACATACACCTACATCTCCAGATTAATTAACATACAC 1791  
 Db 1881 GAACCCCCCGGTTCCCTTCTCGACATACACCTACATCTCCAGATTAATTAACATACAC 1940  
 QY 1792 TGACCTCATCTGCTCCGAGAGACCCATTAAAG-TTGGGTAATTTAAAGTATCGCAATAC 1850  
 Db 1941 TGACCTCATCTGCTCCGAGAGACCCATTAAAGTTGGGTAATTTAAAGTATCGCAATAC 2000  
 QY 1851 CAGATGAGACAAACAGTAACAACTACTGCGCATTTCCCTGAGCTAGCCGAGG 1910  
 Db 2001 CAGATGAGACAAACAGTAACAACTACTGCGCATTTCCCTGAGCTAGCCGAGG 2060  
 QY 1911 AGCAGGCTTTGAGACTCTGTTGGACACAGTTGGTGTGACGTGCGCCAGAGAGCTCT 1970  
 Db 2061 AGCAGGCTTTGAGACTCTGTTGGACACAGTTGGTGTGACGTGCGCCAGAGAGCTCT 2120  
 QY 1971 ACTGACCAACATGACTGCTTCAAGTCTGCTATCAGAGATGCTCTTATCAATTGTGA 2030  
 Db 2121 ACTGACCAACATGACTGCTTCAAGTCTGCTATCAGAGATGCTCTTATCAATTGTGA 2180  
 QY 2031 TGATTTGAACTTTTAAACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2090  
 Db 2181 TGATTTGAACTTTTAAACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2240  
 QY 2091 CCATTGTTTCCAGACAAAGTACTTATTAATTAATTTAGTGTTCGACAGA 2143  
 Db 2241 CCATTGTTTCCAGACAAAGTACTTATTAATTAATTTAGTGTTCGACAGA 2293

RESULT 9  
 AAK52699  
 ID AAK52699 standard; cDNA; 3003 BP.  
 XX  
 AC AAK52699;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 2228.  
 XX  
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukemia;  
 KW nervous system disorder; arthritis; inflammation; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200157190-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PE 05-FEB-2001; 2001WO-US04098.  
 XX  
 PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 PR 19-JUN-2000; 2000US-0598075.  
 PR 19-JUL-2000; 2000US-0620325.  
 PR 01-SEP-2000; 2000US-0654936.  
 PR 15-SEP-2000; 2000US-0683361.  
 PR 20-OCT-2000; 2000US-0693325.  
 PR 30-NOV-2000; 2000US-0728422.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX

PI Tang YF, Liu C, Dymnac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao Qa, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 XX  
 DR WPI: 2001-476283/51.  
 P-PSDB: AAM79566.  
 PT Nucleic acids encoding polypeptides with cytokine-like activities,  
 XX useful in diagnosis and gene therapy -  
 PS Claim 1; Page 4584-4585; 6221pp; English.  
 XX  
 CC The invention relates to polynucleotides (AAK51456-AAK51435) and the  
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAM80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.  
 XX  
 SQ Sequence 3003 BP; 807 A; 703 C; 721 G; 772 T; 0 other;  
 Query Match 93.4%; Score 2008.2; DB 22; Length 3003;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 2024; Conservative 0; Mismatches 8; Indels 1; Gaps 1;  
 QY 112 TTATGAAATTCCTGAGAGAGAGAAAGACTGTAATATGCGGACCCCTAGAGAAATATA 171  
 Db 298 TTCCAGAAATTCCTGAGAGAGAGAAAGACTGTAATATGCGGACCCCTAGAGAAATATA 357  
 QY 172 CCAGAGAAAGAAATTCCTGAGAGAGAAAGACTGTAATATGCGGACCCCTAGAGAAATATA 231  
 Db 358 CCAGAGAAAGAAATTCCTGAGAGAGAAAGACTGTAATATGCGGACCCCTAGAGAAATATA 417  
 QY 232 GAAACATGATGTTTGAAGACACTGCTATGAGACAGTGAATGTTGTGCGCAAAACAAA 291  
 Db 418 GAAACATGATGTTTGAAGACACTGCTATGAGACAGTGAATGTTGTGCGCAAAACAAA 477  
 QY 292 CTGCAATGACACTTCAGATGATGTTGCGCAAGAGAGAAAGAAATCTAGACAACTAT 351  
 Db 478 CTGCAATGACACTTCAGATGATGTTGCGCAAGAGAGAAAGAAATCTAGACAACTAT 537  
 QY 352 GAAAGGAAAGAAAGAAAGTGTGCAAAATCTTGGACAGTGTGCAAGTCAAGTCAAGT 411  
 Db 538 GAAAGGAAAGAAAGAAAGTGTGCAAAATCTTGGACAGTGTGCAAGTCAAGTCAAGT 597  
 QY 412 GAAATGGAACATCTTATATCCCAATGTGCTATCCCAATGAGGACATTAACACTG 471  
 Db 598 GAAATGGAACATCTTATATCCCAATGTGCTATCCCAATGAGGACATTAACACTG 657  
 QY 472 TATCTTAACCTATGTTGAGAGAGATTTGCAATGCTGCTGAGGATGATGAT 531  
 Db 658 TATCTTAACCTATGTTGAGAGAGATTTGCAATGCTGCTGAGGATGATGAT 717  
 QY 532 CATATGCTGAGAAACATCTGCTATACCTGAGATGCCAAATCACTATGCTGTAACCT 591  
 Db 718 CATATGCTGAGAAACATCTGCTATACCTGAGATGCCAAATCACTATGCTGTAACCT 777  
 QY 592 GTGTGCAAGAAATGTACGAGAGTACCTGATGAGGATCTGTGGAAGAAAGCTTATCGAG 651  
 Db 778 GTGTGCAAGAAATGTACGAGAGTACCTGATGAGGATCTGTGGAAGAAAGCTTATCGAG 837  
 QY 652 AGAATGTCAGACAGATTCCTGTGTGAGAGGCTGTGCAAGAAAGAGATGAGGACAG 711  
 Db 838 AGAATGTCAGACAGATTCCTGTGTGAGAGGCTGTGCAAGAAAGAGATGAGGACAG 897

QY 712 TATTATTCAAAAACAACTCTGACGGAGATGCTCTCCCAACTTTTATAGACA 771  
 |||||||  
 Db 898 TATTATTCAAAAACAACTCTGACGGAGATGCTCTCCCAACTTTTATAGACA 957  
 QY 772 CTTTATCTTAAATATATCAAGATGAGACATATAGATATATATGAGATGGAAGA 831  
 |||||||  
 Db 958 CTTTATCTTAAATATATCAAGATGAGACATATAGATATATATGAGATGGAAGA 1017  
 QY 832 CATAGTTACAGAGATTCACCTGCCAAGTGAACAGCAAGAGATTACTGTTTACAG 891  
 |||||||  
 Db 1018 CATAGTTACAGAGATTCACCTGCCAAGTGAACAGCAAGAGATTACTGTTTACAG 1077  
 QY 892 TATGATGATCGAAAAATATGTAAGGCGCTTGAGACAAACAAATCAGATGCGGATAAA 951  
 |||||||  
 Db 1078 TATGATGATCGAAAAATATGTAAGGCGCTTGAGACAAACAAATCAGATGCGGATAAA 1137  
 QY 952 AACACATTTGGAATGCAAGCAATTCCTCAGAGCCATACAGTTGAGTCTGCTCCAG 1011  
 |||||||  
 Db 1138 AACACATTTGGAATGCAAGCAATTCCTCAGAGCCATACAGTTGAGTCTGCTCCAG 1197  
 QY 1012 TATGATGAGAGATGATCATTAACAGATCATCGATTCACAGTCAAGTGGGATGTA 1071  
 |||||||  
 Db 1198 TATGATGAGAGATGATCATTAACAGATCATCGATTCACAGTCAAGTGGGATGTA 1257  
 QY 1072 AATACAGTGAATGCTTAAACAGTGTGATCACCATTGTGAAGCATCTCTGACTTGGCT 1131  
 |||||||  
 Db 1258 AATACAGTGAATGCTTAAACAGTGTGATCACCATTGTGAAGCATCTCTGACTTGGCT 1317  
 QY 1132 TTCAATTAATGCGATGATGCTGACCTGCTCCAAAGATTCCTTCCATGCTGATGAGATATG 1191  
 |||||||  
 Db 1318 TTCAATTAATGCGATGATGCTGACCTGCTCCAAAGATTCCTTCCATGCTGATGAGATATG 1377  
 QY 1192 GCTCTCCCACTGACATTCCTCCGGAGGCTGCTGTCGGACACCGAGCTGCTCAAT 1251  
 |||||||  
 Db 1378 GCTCTCCCACTGACATTCCTCCGGAGGCTGCTGTCGGACACCGAGCTGCTCAAT 1437  
 QY 1252 GTTGTGACCTTTGATGACACATGATGTTTCTGCATCTGGGAGATGAACATTAAGGTA 1311  
 |||||||  
 Db 1438 GTTGTGACCTTTGATGACACATGATGTTTCTGCATCTGGGAGATGAACATTAAGGTA 1497  
 QY 1312 TGGACACAAATGATCTTGTGAATTTGTAGGACCTTAATGACACAAAGAGCATTTGCC 1371  
 |||||||  
 Db 1498 TGGACACAAATGATCTTGTGAATTTGTAGGACCTTAATGACACAAAGAGCATTTGCC 1557  
 QY 1372 TGTTCGATGACAGGAGGAGGCTGAGTGAAGGCTCATCTGACAACTATCAGATTA 1431  
 |||||||  
 Db 1558 TGTTCGATGACAGGAGGAGGCTGAGTGAAGGCTCATCTGACAACTATCAGATTA 1617  
 QY 1432 TGGGACATAGAAATGCTGATGCTTACAGATGTTTGAAGGCCATGAGGAATTTGGTGGT 1491  
 |||||||  
 Db 1618 TGGGACATAGAAATGCTGATGCTTACAGATGTTTGAAGGCCATGAGGAATTTGGTGGT 1677  
 QY 1492 TGTATTCGATTTGATTAACAGAGATATGTCAGTGGGCGCTATGATGGAATAATTAAGTG 1551  
 |||||||  
 Db 1678 TGTATTCGATTTGATTAACAGAGATATGTCAGTGGGCGCTATGATGGAATAATTAAGTG 1737  
 QY 1552 TGGGATCTTGTGCTGCTTGGACCCCGTGTCTGACAGGACATCTGTCTACGGACC 1611  
 |||||||  
 Db 1738 TGGGATCTTGTGCTGCTTGGACCCCGTGTCTGACAGGACATCTGTCTACGGACC 1797  
 QY 1612 CTTGTGAGGACATTCGGAAGATTTTTCGACTACAGTTGATGAATTCAGATTTGCACT 1671  
 |||||||  
 Db 1798 CTTGTGAGGACATTCGGAAGATTTTTCGACTACAGTTGATGAATTCAGATTTGCACT 1857  
 QY 1672 AGTTCAATGATGACAAATCTCATCTGGGACTTCTTAATGATCAGGTGCGCAAGCT 1731  
 |||||||  
 Db 1858 AGTTCAATGATGACAAATCTCATCTGGGACTTCTTAATGATCAGGTGCGCAAGCT 1917  
 QY 1732 GAAACCCCGCTTCCCTTCTGACACATACACCTACATCTCCAGATTAATTAACATACAG 1791  
 |||||||  
 Db 1918 GAAACCCCGCTTCCCTTCTGACACATACACCTACATCTCCAGATTAATTAACATACAG 1977  
 QY 1792 TGACCTCATCTTCCCGACAGACCCCATTAAG-17TGGCGTATTTAACGTAATCTCCCAATAC 1850

Db 1978 TGACCTCATCTTCCCGACAGACCCCATTAAGTTTGGGATTTAAGTATCTGCCAATAC 2037  
 |||||||  
 QY 1851 CAGGATGACCAACACAGTAAATCAATCAATCTGCCCCAGTTTCCCTGACATACCGGAG 1910  
 |||||||  
 Db 2038 CAGGATGACCAACACAGTAAATCAATCAATCTGCCCCAGTTTCCCTGACATACCGGAG 2097  
 QY 1911 AGCAGGCGTTTGGACTCTGTTGGGACACAGTGGTGTGACAGTGGCGCCAGAGGCTCT 1970  
 |||||||  
 Db 2098 AGCAGGCGTTTGGACTCTGTTGGGACACAGTGGTGTGACAGTGGCGCCAGAGGCTCT 2157  
 QY 1971 ACACAGACAACTGACTGCTTCACTGCTCTATCAGAGATGCTTTTATCAATTTGGA 2030  
 |||||||  
 Db 2158 ACACAGACAACTGACTGCTTCACTGCTCTATCAGAGATGCTTTTATCAATTTGGA 2217  
 QY 2031 TGATTGAACTTTTAAACCTCCCTCTCTCTCTCTTCACTGACCTGACCTAGTTTTC 2090  
 |||||||  
 Db 2218 TGATTGAACTTTTAAACCTCCCTCTCTCTCTCTTCACTGACCTGACCTAGTTTTC 2277  
 QY 2091 CCAATGCTTCAGACAAAGGTGACTTATAATTAATTTAGTGTGTTGCCAGAA 2143  
 |||||||  
 Db 2278 CCAATGCTTCAGACAAAGGTGACTTATAATTAATTTAGTGTGTTGCCAGAA 2230  
 |||||||  
 RESULT 10  
 AAK52700  
 .ID AAK52700 standard; cDNA; 3003 BP.  
 XX AC AAK52700;  
 XX DT 06-NOV-2001 (first entry)  
 XX DE Human polynucleotide SEQ ID NO 2229.  
 DE XX  
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation; ss.  
 OS Homo sapiens.  
 XX PN W0200157190-A2.  
 XX PD 09-AUG-2001.  
 XX PF 05-FEB-2001; 2001WO-US04098.  
 XX PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 PR 20-JUN-2000; 2000US-0598075.  
 PR 19-JUL-2000; 2000US-0620325.  
 PR 01-SEP-2000; 2000US-0654936.  
 PR 15-SEP-2000; 2000US-0663361.  
 PR 20-OCT-2000; 2000US-0693325.  
 PR 30-NOV-2000; 2000US-0728422.  
 XX PA (HXSE-) HXSEQ INC.  
 XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wehrman T, Goodrich R;  
 XX DR WPI: 2001-476283/51.  
 DR P-PSDB: AAM79567.  
 XX PT Nucleic acids encoding and gene therapy -  
 PT useful in diagnosis and gene therapy -  
 XX PS Claim 1: Page 4585-4586; 6221pp; English.  
 XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce





```

Db 1318 TTCATTAATGGCAGTATGCTGACCTGCTCAAGATGCTTCATTGCTATGGGATATG 1377
QY 1192 GCTTCCCACTGACATTAACCTCCGAGGGTCTGTGGAGACACGAGCTGTGTCAAT 1251
Db 1378 GCTTCCCACTGACATTAACCTCCGAGGGTCTGTGGAGACACGAGCTGTGTCAAT 1437
QY 1252 GTTGTAGACTTGTAGACAGTACTTGTCTGATCTGGGATAGAACTATTAAGTA 1311
Db 1438 GTTGTAGACTTGTAGACAGTACTTGTCTGATCTGGGATAGAACTATTAAGTA 1497
QY 1312 TGGACACAGTACTTGTGATTTGTAAAGACCTTAAATGACAAACGAGCATTTGCC 1371
Db 1498 TGGACACAGTACTTGTGATTTGTAAAGACCTTAAATGACAAACGAGCATTTGCC 1557
QY 1372 TGTGTGCGACTAGAGGAGCGCTGTAGTGTAGTGGCTCATCTGACACATCATAGATTA 1431
Db 1558 TGTGTGCGACTAGAGGAGCGCTGTAGTGTAGTGGCTCATCTGACACATCATAGATTA 1617
QY 1432 TGGGACATAGATGTGTGCTGCTTACAGTGTAAAGGCGCATGAGGAATTTGGTGGT 1491
Db 1618 TGGGACATAGATGTGTGCTGCTTACAGTGTAAAGGCGCATGAGGAATTTGGTGGT 1677
QY 1492 TGTATTCGATTTGATTAACAGAGATAGTCACTGAGGGCTTATGATGAAAAATTAAGTG 1551
Db 1678 TGTATTCGATTTGATTAACAGAGATAGTCACTGAGGGCTTATGATGAAAAATTAAGTG 1737
QY 1552 TGGGATCTTGTGGCTGCTTTGAGCCCCCGTCTCTCTGAGGAGACACTCTCTACGAGAC 1611
Db 1738 TGGGATCTTGTGGCTGCTTTGAGCCCCCGTCTCTCTGAGGAGACACTCTCTACGAGAC 1797
QY 1612 CTGTGTGAGCATTCGCGAAGAGTTTGTGACTACAGTTTGTGATTTCCAGATGTGACT 1671
Db 1798 CTGTGTGAGCATTCGCGAAGAGTTTGTGACTACAGTTTGTGATTTCCAGATGTGACT 1857
QY 1672 AGTTACATGTATGACACATCTCATCTGAGGACTTCCCTAAATGATTCAGCTGCCAATCT 1731
Db 1858 AGTTACATGTATGACACATCTCATCTGAGGACTTCCCTAAATGATTCAGCTGCCAATCT 1917
QY 1732 GAAACCCCGCTGCTGCTTGTGAAACATACACATCTCTCAATTAATTAATTAATTAAT 1791
Db 1918 GAAACCCCGCTGCTGCTTGTGAAACATACACATCTCTCAATTAATTAATTAATTAAT 1977
QY 1792 TGACCTCATCTTGGCCGAGACCATTAAG-TTGGGATTTTAAGCTATCTGCCAATTA 1850
Db 1978 TGACCTCATCTTGGCCGAGACCATTAAG-TTGGGATTTTAAGCTATCTGCCAATTA 2037
QY 1851 CAGATGAGCAACAGTAACAATCAAACTAGTCCAGTTTCCCTGAGCTAGCCGAGG 1910
Db 2038 CAGATGAGCAACAGTAACAATCAAACTAGTCCAGTTTCCCTGAGCTAGCCGAGG 2097
QY 1911 AGCAGGGCTTGTGAGACCTGCTGTGGAGACAGTGTGCTGTGAGTGGCCGAGAGCGTCT 1970
Db 2098 AGCAGGGCTTGTGAGACCTGCTGTGGAGACAGTGTGCTGTGAGTGGCCGAGAGCGTCT 2157
QY 1971 ACTGAGCAGCAACTGCTGCTGCTGCTGCTATCAGAAAGATGCTTCTATCAATTTGTGA 2030
Db 2158 ACTGAGCAGCAACTGCTGCTGCTGCTGCTATCAGAAAGATGCTTCTATCAATTTGTGA 2217
QY 2031 TGATTTGGAATTTTAAACCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2090
Db 2218 TGATTTGGAATTTTAAACCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2277
QY 2091 CCATTTGCTGAGCAAGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2143
Db 2278 CCATTTGCTGAGCAAGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2330

```

```

RESULT 12
AAA73132
ID AAA73132 standard; cDNA; 1707 BP.
XX
AC AAA73132;
XX

```

```

DT 27-NOV-2000 (first entry)
XX
DE Human beta-transducin repeat containing protein (beta-Trcp) cDNA.
XX
KW Ubiquitin ligase SCF complex; F-box protein; ubiquitination; Ikappab;
KW beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FMD1;
KW gene therapy; colon cancer; beta-transducin repeat containing protein;
KW beta-Trcp; ss.
XX
OS Homo sapiens.
XX
PN JP2000166542-A.
XX
PD 20-JUN-2000.
XX
PF 02-DEC-1998; 98JP-0343437.
XX
PR 02-DEC-1998; 98JP-0343437.
XX
PA (KAGA-) KAGAKU GIYUTSU SHINKO JIGYODAN.
XX
DR WPI; 2000-48550/43.
XX
DR P-PDB; AAB12813.
XX
PT F-box protein of ubiquitin ligase SCF complex which promotes the
PT ubiquitination of Ikappab or beta-catenin
XX
PS Disclosure; Fig 17; 19pp; Japanese.
XX
CC The present invention describes an F-box motif protein of ubiquitin
CC ligase SCF complex which promotes the ubiquitination of Ikappab or
CC beta-catenin and is constituted by Skp1 protein, Cull1 protein and a
CC complex (SCF complex) of F-box protein containing F-box motif and WD40
CC repeat motif and has the amino acid sequence of 45 residues (AAB12811)
CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin
CC ligase FMD1 protein) and (AAB12813, which is human beta-transducin
CC repeat containing protein (beta-Trcp)). The F-box protein can be used for
CC the gene therapy of colon cancer by being recombined to a virus vector.
CC The present sequence encodes the human beta-Trcp protein from the
CC present invention.
XX
SO Sequence 1707 BP; 514 A; 349 C; 414 G; 430 T; 0 other;
XX
Query Match 79.4%; Score 1707; DB 21; Length 1707;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 70 ATGAGCCCGGCGGAGGCGGTGTCGAAGAGAGCACTGAATTTGATTCCTCAGAG 129
Db 1 ATGAGCCCGGCGGAGGCGGTGTCGAAGAGAGCACTGAATTTGATTCCTCAGAG 60
QY 130 AGAAGAGCTGTATTAATGCGGAGAACCCCTAGAGATTAATACAGAGAAATTCACCTT 189
Db 61 AGAAGAGCTGTATTAATGCGGAGAACCCCTAGAGATTAATACAGAGAAATTCACCTT 120
QY 190 AGCAGACATACACAGCTGTGCGAGACTCTGCTTAACCAAGAAACAGTATGTTAGCA 249
Db 121 AGCAGACATACACAGCTGTGCGAGACTCTGCTTAACCAAGAAACAGTATGTTAGCA 180
QY 250 AGCAGCTGTATGAGAGCTGAGAAATTTGTGGCCAAACAAACTTGGCCAAATGGACTTGC 309
Db 181 AGCAGCTGTATGAGAGCTGAGAAATTTGTGGCCAAACAAACTTGGCCAAATGGACTTGC 240
QY 310 AGTATGATTTGGCCCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 369
Db 241 AGTATGATTTGGCCCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 300
QY 370 TGTGTCAAATTAATTTGAGCAGTGTGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 429
Db 301 TGTGTCAAATTAATTTGAGCAGTGTGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 360
QY 430 ATATCCCAATGTGTGATTAACCAACATGGGCACTTAATCTGTATCTTAACCTTATGTTG 489

```

```

Db 361 ATATCCAAATGTGTCATTTACCAACATGGGACATMAACTGTAATCTTAACTATGTTG 420
Oy 490 CAGAGAGATTTTCATTAAGTCTCTGCGACCTGGGAGATTGATCATATGCTGAGAACAT 549
Db 421 CAGAGAGATTTTCATTAAGTCTCTGCGACCTGGGAGATTGATCATATGCTGAGAACAT 480
Oy 550 CTGTATACCTTGGATGCGCAATCTATGCTGCTGTAACCTTGTGTGCAAGAAATGATC 609
Db 481 CTGTATACCTTGGATGCGCAATCTATGCTGCTGTAACCTTGTGTGCAAGAAATGATC 540
Oy 610 CGAGTACCTTGGATGCGCAATCTATGCTGCTGTAACCTTGTGTGCAAGAAATGATC 669
Db 541 CGAGTACCTTGGATGCGCAATCTATGCTGCTGTAACCTTGTGTGCAAGAAATGATC 600
Oy 670 TCTGTGTGAGAGGCGCTGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 729
Db 601 TCTGTGTGAGAGGCGCTGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Oy 730 CCTCTGACGGGAGATGCTCTCCCAACTCTTTTATAGAGCACTTATCTTAAATTTAT 789
Db 661 CCTCTGACGGGAGATGCTCTCCCAACTCTTTTATAGAGCACTTATCTTAAATTTAT 720
Oy 790 CAAGCATTTGAGACATTAATGAAATCTAATGAGATGAGAGATGAGATGAGAGATGAG 849
Db 721 CAAGCATTTGAGACATTAATGAAATCTAATGAGATGAGAGATGAGATGAGAGATGAG 780
Oy 850 CACTGCGGAGTTGAAACAAAGAGATTTACTCTTTTACAGTATGATGATGACAAATA 909
Db 781 CACTGCGGAGTTGAAACAAAGAGATTTACTCTTTTACAGTATGATGATGACAAATA 840
Oy 910 GTAAGGCGCTTGCAGACAAACAAATCAAGATCTGGATTAACAAACATTTGAGATGAG 969
Db 841 GTAAGGCGCTTGCAGACAAACAAATCAAGATCTGGATTAACAAACATTTGAGATGAG 900
Oy 970 CGAATTTTCACAGGCGCATACAGGTTCACTCTCTCTCCAGTATGATGACAGATGATC 1029
Db 901 CGAATTTTCACAGGCGCATACAGGTTCACTCTCTCTCCAGTATGATGATGACAGATGATC 960
Oy 1030 ATAACAGATTCATCGGATTCACAGGTTCACTCTCTCTCCAGTATGATGATGACAGATGATC 1089
Db 961 ATAACAGATTCATCGGATTCACAGGTTCACTCTCTCTCCAGTATGATGATGACAGATGATC 1020
Oy 1090 AACAGTTCATTCACAGGTTCACTCTCTCTCTCCAGTATGATGATGACAGATGATC 1149
Db 1021 AACAGTTCATTCACAGGTTCACTCTCTCTCTCCAGTATGATGATGACAGATGATC 1080
Oy 1150 GTGACCTGCTCCAAAGATGCTTCATTTGCTGATGAGATGAGCCCTCCCAACTGATC 1209
Db 1081 GTGACCTGCTCCAAAGATGCTTCATTTGCTGATGAGATGAGCCCTCCCAACTGATC 1140
Oy 1210 ACCCTGCGGAGGCTGCTGCTGCGACACGAGTGCATGCTGATGAGATGAGATGATC 1269
Db 1141 ACCCTGCGGAGGCTGCTGCTGCGACACGAGTGCATGCTGATGAGATGAGATGATC 1200
Oy 1270 AAGTACATTTGCTTCTGATCTGCGGATGAGATTAAGGATGAGACCAAGTACTGTT 1329
Db 1201 AAGTACATTTGCTTCTGATCTGCGGATGAGATTAAGGATGAGACCAAGTACTGTT 1260
Oy 1330 GAATTTGTAAGGACCTTAAATGAGACCAAGGAGATGCTGTTTGGAGTACAGGAGC 1389
Db 1261 GAATTTGTAAGGACCTTAAATGAGACCAAGGAGATGCTGTTTGGAGTACAGGAGC 1320
Oy 1390 AGGCTGTGATGAGGCTGATCTGACAACTATGATGAGTATGAGATGAGATGATGAT 1449
Db 1321 AGGCTGTGATGAGGCTGATCTGACAACTATGATGAGTATGAGATGAGATGATGAT 1380
Oy 1450 GCATTTGTAAGGATTTTAAAGGACCTGAGATGAGTCTGTTTGGAGTACAGGAGC 1509
Db 1381 GCATTTGTAAGGATTTTAAAGGACCTGAGATGAGTCTGTTTGGAGTACAGGAGC 1440
Oy 1510 AAGAGATGATGATGAGGCGCTTATGAGTAAAGTGTGGAGTCTTGGCTGCT 1569
Db 1441 AAGAGATGATGATGAGGCGCTTATGAGTAAAGTGTGGAGTCTTGGCTGCT 1500

```

```

Oy 1570 TTGAGACCCCGTCTCTCTGAGGAGACATCTGTCTACGAGACCTTGTGAGCATTCGGA 1629
Db 1501 TTGAGACCCCGTCTCTCTGAGGAGACATCTGTCTACGAGACCTTGTGAGCATTCGGA 1560
Oy 1630 AGAGTTTTCGACATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1689
Db 1561 AGAGTTTTCGACATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
Oy 1690 ATCTCATCTGAGGACCTCTCTAAATGATCCAGTCCCAAGCTGAGACCCCGTCCCT 1749
Db 1621 ATCTCATCTGAGGACCTCTCTAAATGATCCAGTCCCAAGCTGAGACCCCGTCCCT 1680
Oy 1750 TCTGAGACATACACCTCATCTCCAGA 1776
Db 1681 TCTGAGACATACACCTCATCTCCAGA 1707

RESULT 13
AAK51715
ID AAK51715 standard; cDNA; 2285 BP.
XX
AC AAK51715;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 260.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
FN W0200157190-A2.
XX
PD 09-AUG-2001.
XX
PE 05-FEB-2001; 2001MO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YF, Liu C, Dymnac RF, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AD, Yang Y, Wejrtman T, Goodrich R;
XX
WP1: 2001-476283/51.
DR P-PSDB; AAM78582.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 1; Page 1177-1180; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and/or
CC activin/inhibin activity and may be useful in the diagnosis and/or

```



PN W0200157190-A2.  
XX  
XX 09-AUG-2001.  
PD  
XX  
PE 05-FEB-2001; 2001MO-US04098.  
XX  
PR 03-FEB-2000; 2000US-0469614.  
PR 27-APR-2000; 2000US-0560875.  
PR 20-JUN-2000; 2000US-0598075.  
PR 19-JUL-2000; 2000US-0620325.  
PR 01-SEP-2000; 2000US-0624936.  
PR 15-SEP-2000; 2000US-0663561.  
PR 20-OCT-2000; 2000US-0693325.  
PR 30-NOV-2000; 2000US-0728422.  
XX  
XX (HXSE-) HXSEQ INC.  
XX  
XX Tang Y<sup>T</sup>, Liu C, Drmanac R<sup>T</sup>, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
PI Zhao Q<sup>A</sup>, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
XX  
XX WPI: 2001-476283/51.  
DR P-PSDB: AAM78584.  
XX  
XX Nucleic acids encoding polypeptides with cytokine-like activities,  
PT useful in diagnosis and gene therapy -  
XX  
XX Claim 1; Page 1183-1186; 6221pp; English.  
XX  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78333-AAK80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
CC (AAM80020) are omitted as the relevant pages from the sequence listing  
CC were missing at the time of publication.  
XX  
XX Sequence 2366 BP; 661 A; 519 C; 594 G; 592 T; 0 other;  
SQ  
Query Match 78.9%; Score 1697.8; DB 22; Length 2366;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 1708; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
QY 112 TTTATGAATTCCTCAGAGAGAGAAAGACTGTAATAATGAGCGAACCCTTAGAGAGATAATA 171  
DB 358 TTCAATGAATTCCTCAGAGAGAGAAAGACTGTAATAATGAGCGAACCCTTAGAGAGATAATA 417  
QY 172 CCAGAGAGAAATTCCTCAGAGAGAGAAAGACTGTAATAATGAGCGAACCCTTAGAGAGATAATA 231  
DB 418 CCAGAGAGAAATTCCTCAGAGAGAGAAAGACTGTAATAATGAGCGAACCCTTAGAGAGATAATA 477  
QY 232 GAAACAGATATGTTAGCAAGCACTGCTATGAAGACTGGAATTTGTGGCCAAACAAAA 291  
DB 478 GAAACAGATATGTTAGCAAGCACTGCTATGAAGACTGGAATTTGTGGCCAAACAAAA 537  
QY 292 CTGGCCAAATGGCACTTCAGATGATGTTGTGCCCAAGCAAGGAAATCTCTGCAAGCTAT 351  
DB 538 CTGGCCAAATGGCACTTCAGATGATGTTGTGCCCAAGCAAGGAAATCTCTGCAAGCTAT 597  
QY 352 GAAAGAGAAAGAAAGAACTGTGTCAAAATCTTTGAGAGAGTGTAGAGTGCATCAAGTG 411  
DB 598 GAAAGAGAAAGAAAGAACTGTGTCAAAATCTTTGAGAGAGTGTAGAGTGCATCAAGTG 657  
QY 412 GAATTTGTGGAACATCTTATATCCCAATGTGTGATTACCAACATGGGCACATAACTCG 471  
DB 658 GAATTTGTGGAACATCTTATATCCCAATGTGTGATTACCAACATGGGCACATAACTCG 717

QY 472 TATCTTAACCTATGTTGACAGAGAGATTTCAATACCTGCTCCAGCTCGGGGATTTGAT 531  
DB 718 TATCTTAACCTATGTTGACAGAGAGATTTCAATACCTGCTCCAGCTCGGGGATTTGAT 777  
QY 532 CATATGCTGAGAACTTGTGTCATACCTGAGATGCAATTCATATGTCGCTGAACCT 591  
DB 778 CATATGCTGAGAACTTGTGTCATACCTGAGATGCAATTCATATGTCGCTGAACCT 837  
QY 592 GTGTGCAAGAAATGATACAGAGTACCTCTGTATGTCATGCTGTGGAAGACTTATCAG 651  
DB 838 GTGTGCAAGAAATGATACAGAGTACCTCTGTATGTCATGCTGTGGAAGACTTATCAG 897  
QY 652 AGAATGCTGAGAGACATTTCTGTCGAGAGAGGCTCGGCAAGCAAGAGATGGGGACG 711  
DB 898 AGAATGCTGAGAGACATTTCTGTCGAGAGAGGCTCGGCAAGCAAGAGATGGGGACG 957  
QY 712 TATTTATCAAAAACAAACCTCTGACGGGAATGCTCCCACTCTTTTATAGACA 771  
DB 958 TATTTATCAAAAACAAACCTCTGACGGGAATGCTCCCACTCTTTTATAGACA 1017  
QY 772 CTTTATCTTAATTTATACAGACATTTGACATATGAATCTAATTTGAGATGTGAAGA 831  
DB 1018 CTTTATCTTAATTTATACAGACATTTGACATATGAATCTAATTTGAGATGTGAAGA 1077  
QY 832 CATAGTTTACAGAGATTTGACATGCGGAGTGAACAGCAAGAGATTTACTGTTACAG 891  
DB 1078 CATAGTTTACAGAGATTTGACATGCGGAGTGAACAGCAAGAGATTTACTGTTACAG 1137  
QY 892 TATGATGATCAGAAATTAAGTAAAGCGGCTTGAGACACACATCAAGATCTGGGATAA 951  
DB 1138 TATGATGATCAGAAATTAAGTAAAGCGGCTTGAGACACACATCAAGATCTGGGATAA 1197  
QY 952 AATACATTTGAATGCAAGGAATTTTCACAGGCAATACAGTTCAGTCTCTGTCGAG 1011  
DB 1198 AATACATTTGAATGCAAGGAATTTTCACAGGCAATACAGTTCAGTCTCTGTCGAG 1257  
QY 1012 TATGATGAGAGATGATCAATACAGAGATCATGATTCACGCGTGCAGAGCTGTGATGA 1071  
DB 1258 TATGATGAGAGATGATCAATACAGAGATCATGATTCACGCGTGCAGAGCTGTGATGA 1317  
QY 1072 AATACAGGCAATTTGCAATACAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1131  
DB 1318 AATACAGGCAATTTGCAATACAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1377  
QY 1132 TTCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1191  
DB 1378 TTCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1437  
QY 1192 GCTTCCCAACTGACATTTACCTCCGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1251  
DB 1438 GCTTCCCAACTGACATTTACCTCCGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1497  
QY 1252 GTTGTGACATTTGATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1311  
DB 1498 GTTGTGACATTTGATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1557  
QY 1312 TGGACACAAAGTACTTGTGAATTTGTGAAGACCTTAAATGACACAAAGAGCAATGCG 1371  
DB 1558 TGGACACAAAGTACTTGTGAATTTGTGAAGACCTTAAATGACACAAAGAGCAATGCG 1617  
QY 1372 TGTTTGACAGACAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1431  
DB 1618 TGTTTGACAGACAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1677  
QY 1432 TGGACATGAATTTGTGTGATTTTACAGATGTTTGAAGGCGCATGAGAAATTTGGTGGT 1491  
DB 1678 TGGACATGAATTTGTGTGATTTTACAGATGTTTGAAGGCGCATGAGAAATTTGGTGGT 1737  
QY 1492 TGTATTGATTTGATTAACAAGAGATGTCAGTGGGCTTATGATGAGAAAAATTAAGTG 1551  
DB 1738 TGTATTGATTTGATTAACAAGAGATGTCAGTGGGCTTATGATGAGAAAAATTAAGTG 1797

1552 TGGATCTGNGGCTGTTGAGACCCCGCTCTCTGAGGACACTCTGTACGAGAC 1611  
1798 TGGATCTGNGGCTGTTGAGACCCCGCTCTCTGAGGACACTCTGTACGAGAC 1857  
1612 CTGTGAGACATTCGAGAGAGTTTTCGACTACAGTTTGAATCCAGATTGTCAGT 1671  
1858 CTGTGAGACATTCGAGAGAGTTTTCGACTACAGTTTGAATCCAGATTGTCAGT 1917  
1672 AGTTACATGATGACACAACTCTCATCTGGACTTCTTAATGATCCAGCTGCCAAGCT 1731  
1918 AGTTACATGATGACACAACTCTCATCTGGACTTCTTAATGATCCAGCTGCCAAGCT 1977  
1732 GAACCCCCCGCTCTCTGAGACATTCGACTACAGTTTGAATCCAGATTGTCAGT 1791  
1978 GAACCCCCCGCTCTCTGAGACATTCGACTACAGTTTGAATCCAGATTGTCAGT 2037  
1792 TGACCTCATCTGCGCCAGGACCATTAAGTTGCGGTATTAAAC 1836  
2038 TGACCTCATCTGCGCCAGGACCATTAAGTTGCGGTATTAAAC 2082

RESULT 15

AKS1716  
ID AKS1716 standard; cDNA; 2207 BP.

AKS1716;  
XX

06-NOV-2001 (first entry)  
XX

Human polynucleotide SEQ ID NO 261.  
XX

Human: cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation; ss.

Homo sapiens.  
XX

WO200157190-A2.  
XX

09-AUG-2001.  
XX

05-FEB-2001; 2001WO-US04098.  
XX

03-FEB-2000; 2000US-0496914.  
XX

27-APR-2000; 2000US-0560875.  
XX

20-JUN-2000; 2000US-0598075.  
XX

19-JUL-2000; 2000US-0620325.  
XX

01-SEP-2000; 2000US-0654936.  
XX

15-SEP-2000; 2000US-0663561.  
XX

20-OCT-2000; 2000US-0693325.  
XX

30-NOV-2000; 2000US-0728422.  
XX

(HYSF-) HYSF INC.  
XX

Tang YF, Liu C, Dimaac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
PI Zhao Qa, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
XX

WPI: 2001-476283/51.  
XX

P-PSDB: AAM78583.  
XX

Nucleic acids encoding polypeptides with cytokine-like activities,  
PT useful in diagnosis and gene therapy -  
XX  
Claim 1; Page 1180-1183; 6221pp; English.  
XX  
The invention relates to polynucleotides (AKS1456-AKS3435) and the  
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activity/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: Records for SEQ ID NO 2110 (AKS2581), 2111 (AKS2582) and 3666  
CC (AAM80020) are omitted as the relevant pages from the sequence listing  
CC were missing at the time of publication.

Sequence 2207 BP; 615 A; 478 C; 555 G; 559 T; 0 other:

Query Match 78.5%; Score 1688.8; DB 22; Length 2207;  
Best Local Similarity 94.8%; Pred. No. 0;  
Matches 1769; Conservative 0; Mismatches 67; Indels 30; Gaps 1;

QY 1 TGGTGTGGCTGGGCTGCGACCAAAAGGGGCGCCCGGAGAGCCAGTGGCC 60  
Db 58 TGGTGTGGCTGGGCTGCGACCAAAAGGGGCGCCCGGAGAGCCAGTGGCC 117  
QY 61 TGGGCGATTATGAGCCCGCGAGCGGCTGCTCAAGAGAAAGCACTCAAGTTATGAT 120  
Db 118 TGGGCGATTATGAGCCCGCGAGCGGCTGCTCAAGAGAAAGCACTCAAGTTATGAT 177  
QY 121 TCTCAGAGAGAAAGCACTGTAATTAATGCGAACCCTTAGAGATTAATCCAGAGAG 180  
Db 178 TCTATGCGCCAGGCTCTGTGGCTGCTCCAGCTGCGCGAGAGATGCGCTTCGCTG 237  
QY 181 AATT-----CACTTAGACAGACTATCAACAGCTGT 210  
Db 238 CGATGCTGTATTAACCCAGGAGCTGGCCACTCAAGCTTTCACAGATTAACAGCTGT 297  
QY 211 GCCAGACTCTGCTTAACCAAGAAAGCAAGTATGTTAGCAAGACAGCTATGAAGACTAG 270  
Db 298 GCCAGACTCTGCTTAACCAAGAAAGCAAGTATGTTAGCAAGACAGCTATGAAGACTAG 357  
QY 271 AATTGTGTGGCCAAACAACTTCCCAATGGCACTTCCAGATGATTTGCCCAAGCAA 330  
Db 358 AATTGTGTGGCCAAACAACTTCCCAATGGCACTTCCAGATGATTTGCCCAAGCAA 417  
QY 331 CGGAACCTCTGCAAGCTATGAAAGAAAGAAAGCACTGTGTGCAATTAATTTGAGAG 390  
Db 418 CGGAACCTCTGCAAGCTATGAAAGAAAGAAAGCACTGTGTGCAATTAATTTGAGAG 477  
QY 391 TGGTATGAGTATGATCAAGTGAATTTGTGAACATCTTATATCCCAATGTGCTATAC 450  
Db 478 TGGTATGAGTATGATCAAGTGAATTTGTGAACATCTTATATCCCAATGTGCTATAC 537  
QY 451 CAACATGGGGCACTAATCTGATCTTAAACCTATGTTGCAAGAGATTTGATTAAGTCT 510  
Db 538 CAACATGGGGCACTAATCTGATCTTAAACCTATGTTGCAAGAGATTTGATTAAGTCT 597  
QY 511 CTGCCAGCTGGGGGATTTGATCATATGCTGAGAATCTTGTGATCACTGATGAGCCAAA 570  
Db 598 CTGCCAGCTGGGGGATTTGATCATATGCTGAGAATCTTGTGATCACTGATGAGCCAAA 657  
QY 571 TCACATATGCTGCTGCAACTTGTGCAAGAGATTTGATCAAGTGAAGTCTGATGAGCATG 630  
Db 658 TCACATATGCTGCTGCAACTTGTGCAAGAGATTTGATCAAGTGAAGTCTGATGAGCATG 717  
QY 631 CTGTGGAAGAGCTTATGAGAGAAATGTCAGACAGATTTCTGTGAGAGAGGCTTGCA 690  
Db 718 CTGTGGAAGAGCTTATGAGAGAAATGTCAGACAGATTTCTGTGAGAGAGGCTTGCA 777  
QY 691 GAAGGAAGAGATGGGAGCACTATTTATCAAAACAACTCTGACGAGGAATGCTCT 750  
Db 778 GAAGGAAGAGATGGGAGCACTATTTATCAAAACAACTCTGACGAGGAATGCTCT 837  
QY 751 CCCAAGCTTTTATGAGAGCACTTATCTTAAATTAATTAACAGACATTTGAGACAAATGAA 810  
Db 838 CCCAAGCTTTTATGAGAGCACTTATCTTAAATTAATTAACAGACATTTGAGACAAATGAA 897  
QY 811 TCTAATGAGAGATGGAAGATAGTTTACAGAGATTTACATGCGGAAGTGAACAGC 870



GenCore version 5.1.4.P5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2003, 22:02:00 : Search time 2778 Seconds

(without alignments)  
12540.134 Million cell updates/sec

Title: US-09-601-168b-1

Perfect score: 2151

Sequence: 1 tgcgttgctgcgcctggc.....gtttgcagaaaaaaaaa 2151

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlinu:\*  
5: em\_estlov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estlinu:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrl:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	650.6	30.2	709	12	BE782628 601465751
2	638	29.7	778	14	BE444188
3	635.2	29.5	652	9	AL044095
4	625	29.1	851	13	AL044095 DKFZP434M
5	624.8	29.0	973	14	BM425358 AGENCOURT
6	613.2	28.5	758	10	BM465280 AGENCOURT
					BE296484 601174756

7	608.8	28.3	692	10	BE269297
8	589.4	27.4	828	13	BE822845
9	586.6	27.3	1147	14	BQ948186
10	580.6	27.0	743	12	BG829037
11	566.4	26.3	752	12	BG722472
12	564.4	24.8	789	12	BG518761
13	534.4	24.8	637	13	BE103092
14	529.8	24.6	544	9	AL041438
15	528.4	24.6	888	13	BE455545
16	516.2	24.0	810	14	BM944304
17	500.4	23.3	646	12	BG082149
18	497	23.1	850	14	BQ771102
19	492.2	22.9	541	12	BG058605
20	488.8	22.7	590	10	BE762522
21	483.8	22.5	521	12	BE76123
22	483.8	22.5	945	9	AL556068
23	483.6	22.5	580	12	BG277740
24	472.8	22.0	732	13	BE088274
25	471.4	21.9	658	12	BE781002
26	468	21.8	656	9	AL707470
27	464.8	21.6	972	14	BQ940482
28	459.2	21.3	665	9	AL642623
29	458.8	21.3	704	13	BQ48898
30	456.6	21.2	695	13	BM214068
31	452.8	21.1	682	10	BE109603
32	445.6	20.7	464	9	AL073856
33	445.2	20.7	666	14	BQ211349
34	444.4	20.7	553	9	AA478504
35	443.6	20.6	717	10	AV716543
36	440	20.5	577	10	BE284066
37	438.8	20.4	621	14	BQ383756
38	438.2	20.4	506	9	AL430490
39	425.6	19.8	509	14	BQ340149
40	419.6	19.5	664	13	BM211073
41	417.6	19.4	642	13	BE1278543
42	416.2	19.3	656	13	BM214360
43	414.8	19.3	1081	13	BM542586
44	405.4	18.8	524	9	AA035076
45	404.2	18.8	538	9	AA230588

## ALIGNMENTS

RESULT 1  
LOCUS BE782628 709 bp mRNA linear EST 20-OCT-2000  
DEFINITION 601465751P1 NIH\_MGC\_67 Homo sapiens cDNA clone IMAGE:3868812 5',  
mRNA sequence.  
BE782628  
VERSION BE782628.1 GI:10203826  
KEYWORDS  
SOURCE EST.  
ORGANISM human.  
Homo sapiens

REFERENCE 1 (bases 1 to 709)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: L1AM9617 row: f column: 13  
High quality sequence stop: 655.  
location/Qualifiers  
I..709

FEATURES  
SOURCE

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3868812"
/clone_id="NIH_MGC.67"
/lissue_type="retinoblastoma"
/label="DH10B (phage-resistant)"
/Note="Organ: eye; Vector: PCMV-SFOR6; Site: 1; NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."

BASE COUNT      193 a      156 c      175 g      185 t
ORIGIN

Query Match      30.2%; Score 650.6; DB 12; Length 709;
Best Local Similarity 98.2%; Pred. No. 7.5e-184;
Matches 700; Conservative 0; Mismatches 9; Indels 4; Gaps 4;

QY 1222 GTGCTGTCGACACCGAGCTGCTCAATGTTGTAGACTTGTGATGACAGTACATTTGT 1281
    |||||||
Db 1 GTGCTGTCGACACCGAGCTGCTCAATGTTGTAGACTTGTGATGACAGTACATTTGT 59

QY 1282 TCTGCATCTGGGAGATAGACTATTAAGGTATGAAACACAGTACTTGTGAATTTGTAAAG 1341
    |||||||
Db 60 TCTGCATCTGGGAGATAGACTATTAAGGTATGAAACACAGTACTTGTGAATTTGTAAAG 119

QY 1342 ACCCTAAATGACACAAACGAGGATGCTGTTTCAGTACAGGACAGCTGTAGTG 1401
    |||||||
Db 120 ACCCTAAATGACACAAACGAGGATGCTGTTTCAGTACAGGACAGCTGTAGTG 179

QY 1402 AGTGGCTCATGTGACACACTATCAGATTATGAGACATAGATGTGTGATTTTACGA 1461
    |||||||
Db 180 AGTGGCTCATGTGACACACTATCAGATTATGAGACATAGATGTGTGATTTTACGA 239

QY 1462 GTGTTAGAGGCGCATGAGATTTGGTGGTGTATGATTTGATTAACAGAGATATGTC 1521
    |||||||
Db 240 GTGTTAGAGGCGCATGAGATTTGGTGGTGTATGATTTGATTAACAGAGATATGTC 299

QY 1522 AGTGGGCGCTATGATGAAAAATTAAGTGTGGATCTTGTGGCTTTGGACCCCGT 1581
    |||||||
Db 300 AGTGGGCGCTATGATGAAAAATTAAGTGTGGATCTTGTGGCTTTGGACCCCGT 359

QY 1582 GCTCTCGAGGAGCAGCTGTGCTACGACGACCTTTGTGAGCACTTCCGAGAGCTTTTGA 1641
    |||||||
Db 360 GCTCTCGAGGAGCAGCTGTGCTACGACGACCTTTGTGAGCACTTCCGAGAGCTTTTGA 419

QY 1642 CTACAGTTGATGATTCAGATTTGTCAGTATGTCATGATGACACATCTCATCTGG 1701
    |||||||
Db 420 CTACAGTTGATGATTCAGATTTGTCAGTATGTCATGATGACACATCTCATCTGG 479

QY 1702 GACTTCTTAAATGATCCAGCTGCCCCAAGCTGACCCCCCGTTCCTTCGAAACATAC 1761
    |||||||
Db 480 GACTTCTTAAATGATCCAGCTGCCCCAAGCTGACCCCCCGTTCCTTCGAAACATAC 538

QY 1762 ACCATCATCTCCAGATTAATTAACCATACAGTACCTCATATTTCGCCAGGCCCATTTAAA 1821
    |||||||
Db 539 ACCATCATCTCCAGATTAATTAACCATACAGTACCTCATATTTCGCCAGGCCCATTTAAA 598

QY 1822 GTTGGGATTTTAACGATATCTGCCAATACAGATGAGCAACAACAGTAAACATCAACT 1881
    |||||||
Db 599 GTTGGGATTTTAACGATATCTGCCAATACAGATGAGCAACAACAGTAAACATCAACT 656

QY 1882 ACTGGCCGCTTCCCTGACACTAGCCGAGGAGACAGGCGCTTTGAGACGCTCTGTG 1934
    |||||||
Db 657 ACTGGCCGCTTCCCTGACACTAGCCGAGGAGACAGGCGCTTTGAGACGCTCTGTG 709

RESULT 2
LOCUS      B0444188      778 bp      mRNA
DEFINITION UI-M-EX0-bxj-1-06-0-UI.r1 NIH_BMAP_EX0 Mus musculus cDNA clone
IMAGE:5709077 5', mRNA sequence.
ACCESSION  B0444188
VERSION     B0444188.1 GI:21247300

```

```

KEYWORDS      EST,
SOURCE        house mouse,
ORGANISM      Mus musculus
REFERENCE     NIH-MGC http://mgi.nci.nih.gov/
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE         Unpublished (1999)
JOURNAL       Contact: Robert Strausberg, Ph.D.
COMMENT       Email: cgaabs-remail.nih.gov
              Tissue Procurement: Dr. James Lin, University of Iowa
              cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
              DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
              Clone Distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/MLN at:
              http://image.jiml.gov
              This clone was contributed by the Brain Molecular Anatomy Project
              (BMAP)

FEATURES
source
Location/Qualifiers
1..778
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5709077"
/clone_id="NIH_BMAP_EX0"
/lissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/label="DH10B (T1 phage resistant)"
/Note="Organ: brain; Vector: pTX-asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pTX-asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTGCGTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

BASE COUNT    223 a      181 c      204 g      169 t      1 others
ORIGIN

Query Match      29.7%; Score 638; DB 14; Length 778;
Best Local Similarity 89.5%; Pred. No. 4.9e-180;
Matches 697; Conservative 0; Mismatches 81; Indels 1; Gaps 1;

QY 256 GCTATTAAGCATGAGATTTGTGTGGCCAAACAACTTCCCATGCGCTTCCAGTATG 315
    |||||||
Db 1 GCTATTAAGCATGAGATTTGTGTGGCCAAACAACTTCCCATGCGCTTCCAGTATG 60

QY 316 ATGTGCCAAGAGAGGAACTCTCAGACAGCTATGAAAGAAAGAAAGAACTGTGTGC 375
    |||||||
Db 61 ATGTGCCAAGAGAGGAACTCTCAGACAGCTATGAAAGAAAGAAAGAAAGAGCTGTGTGC 120

QY 376 AATATCTTGGAGCAGTGTGCAAGTCAAGTCAAGTGTGGAATTTGTGGAACATTTATCC 435
    |||||||
Db 121 AATATCTTGGAGCAGTGTGCAAGTCAAGTCAAGTGTGGAATTTGTGGAACATTTATCC 180

QY 436 CAAATGTGCATTACCAACATGAGGACATTAACCTGATCTTAACCTATGTGACAGGA 495
    |||||||
Db 181 CAAATGTGCATTACCAACATGAGGACATTAACCTGATCTTAACCTATGTGACAGGA 240

QY 496 GATTTCATACGCTGCTGCGACGCGGAGTGTGATCATATGCTGAGAACATTTGTGCA 555
    |||||||
Db 241 GATTTCATACGCTGCTGCGACGCGGAGTGTGATCATATGCTGAGAACATTTGTGCA 300

```



DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLM13496 row: 1 column: 15  
 High quality sequence stop: 697.

## FEATURES

## SOURCE

Location/Qualifiers  
 1. 851  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6154838"  
 /clone.lib="NIH\_MGC\_67"  
 /tissue\_type="retinoblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: eye; Vector: PCMV-SORT6; Site: 1; Note:  
 Site: 2; Salt: Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.75 kb. Library constructed by Life  
 Technologies."

BASE COUNT 243 a 196 c 223 g 188 t 1 others

## ORIGIN

## Query Match

Best Local Similarity 29.1%; Score 625; DB 14; Length 851;  
 Best Local Similarity 99.1%; Pred. No. 4.2e-176;  
 Matches 639; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

## ORIGIN

112 TTTATGAAATTCCTCGACAGAGAAAGCTGTATATATGGCGAAGCCCTAGGAGATATA 171  
 Db 207 TTCCAGAAATTCCTCGACAGAGAAAGCTGTATATATGGCGAAGCCCTAGGAGATATA 266  
 172 CCAGAGAAATTCCTCGACAGAGAAAGCTGTATATATGGCGAAGCCCTAGGAGATATA 231  
 Db 267 CCAGAGAAATTCCTCGACAGAGAAAGCTGTATATATGGCGAAGCCCTAGGAGATATA 326  
 233 GAAAGAAATTCCTCGACAGAGAAAGCTGTATATATGGCGAAGCCCTAGGAGATATA 291  
 Db 327 GAAAGAAATTCCTCGACAGAGAAAGCTGTATATATGGCGAAGCCCTAGGAGATATA 386  
 292 CTGGCAATGCGACTTCAGATATGATGTGGCCAGCAAGCAAGCAAGCAAGCAAGCAAG 351  
 Db 387 CTGGCAATGCGACTTCAGATATGATGTGGCCAGCAAGCAAGCAAGCAAGCAAGCAAG 446  
 352 GAAAGAAATTCCTCGACAGAGAAAGCTGTATATATGGCGAAGCCCTAGGAGATATA 411  
 Db 447 GAAAGAAATTCCTCGACAGAGAAAGCTGTATATATGGCGAAGCCCTAGGAGATATA 506  
 412 GAAATTCCTCGACAGAGAAAGCTGTATATATGGCGAAGCCCTAGGAGATATA 471  
 Db 507 GAAATTCCTCGACAGAGAAAGCTGTATATATGGCGAAGCCCTAGGAGATATA 566  
 472 TATCTTAAACCTATGTCGAGAGAAATTCCTCGACAGAGAAAGCTGTATATATGGCGA 531  
 Db 567 TATCTTAAACCTATGTCGAGAGAAATTCCTCGACAGAGAAAGCTGTATATATGGCGA 626  
 532 CATATGCGAGAGAAATTCCTCGACAGAGAAATTCCTCGACAGAGAAATTCCTCGACAG 591  
 Db 627 CATATGCGAGAGAAATTCCTCGACAGAGAAATTCCTCGACAGAGAAATTCCTCGACAG 686  
 592 GTGTGCAAGAAATTCCTCGACAGAGAAATTCCTCGACAGAGAAATTCCTCGACAGAG 651  
 Db 687 GTGTGCAAGAAATTCCTCGACAGAGAAATTCCTCGACAGAGAAATTCCTCGACAGAG 746  
 652 AAGATGTCAGAGAAATTCCTCGACAGAGAAATTCCTCGACAGAGAAATTCCTCGACAGAG 710  
 Db 747 AAGATGTCAGAGAAATTCCTCGACAGAGAAATTCCTCGACAGAGAAATTCCTCGACAGAG 806  
 711 GATTTTAAACCAAACTCTCGACAGAGAAATTCCTCGACAGAGAAATTCCTCGACAGAG 755  
 Db 807 GATTTTAAACCAAACTCTCGACAGAGAAATTCCTCGACAGAGAAATTCCTCGACAGAG 851

RESULT 5  
 BM465280 973 bp mRNA Linear EST 05-FEB-2002  
 LOCUS BM465280 973 bp mRNA Linear EST 05-FEB-2002

DEFINITION AGENCOURT\_6427660 NIH\_MGC\_67 Homo sapiens CDNA clone IMAGE:5502075  
 5', mRNA sequence.  
 ACCESSION BM465280  
 VERSION BM465280.1 GI:18514312  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 973)  
 NIH-MGC <http://imgc.nhl.nih.gov/>.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: c9apds-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLM12139 row: k column: 04  
 High quality sequence stop: 601.

## FEATURES

## SOURCE

Location/Qualifiers  
 1. 973  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5502075"  
 /clone.lib="NIH\_MGC\_67"  
 /tissue\_type="retinoblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: eye; Vector: PCMV-SORT6; Site: 1; Note:  
 Site: 2; Salt: Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.75 kb. Library constructed by Life  
 Technologies."

BASE COUNT 282 a 200 c 245 g 246 t

## ORIGIN

Query Match 29.0%; Score 624.8; DB 13; Length 973;  
 Best Local Similarity 97.2%; Pred. No. 5.2e-176;  
 Matches 690; Conservative 0; Mismatches 12; Indels 8; Gaps 5;  
 658 GTCAGAGAGATTCCTGTCGAGAGAGGCTGCGAAGCAAGAGAGAGAGAGATTTTAA 717  
 Db 20 GTCGCGAGAGATTCCTGTCGAGAGAGGCTGCGAAGCAAGAGAGAGAGAGATTTTAA 79  
 718 TTCAAAAACCAAACTCTCGACAGAGAAATTCCTCGACAGAGAAATTCCTCGACAGAG 777  
 Db 80 TTCAAAAACCAAACTCTCGACAGAGAAATTCCTCGACAGAGAAATTCCTCGACAGAG 139  
 778 CTTAAATTTATACAGAGATTCGAGAGAAATTCCTCGACAGAGAAATTCCTCGACAGAG 837  
 Db 140 CTTAAATTTATACAGAGATTCGAGAGAAATTCCTCGACAGAGAAATTCCTCGACAGAG 199  
 838 TTACAGAGAAATTCGAGAGAAATTCCTCGACAGAGAAATTCCTCGACAGAGAAATTC 897  
 Db 200 TTACAGAGAAATTCGAGAGAAATTCCTCGACAGAGAAATTCCTCGACAGAGAAATTC 259  
 898 GATCAGAAATTTAGTAAGAGGCGCTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 957  
 Db 260 GATCAGAAATTTAGTAAGAGGCGCTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 319  
 958 TTGGAATGCAAGCAATTCGAG 1017  
 Db 320 TTGGAATGCAAGCAATTCGAG 379  
 1018 GAG 1077  
 Db 380 GAG 439  
 1078 GGTGAATGCTAAACAGTTCGATTCACATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1137

Db 440 GGTGAATGCTAAACAGCTTGATTCACCATTTGTAAGCAGATTCTGCACCTTGCTTCAAT 499

QY 1138 AATGGCATGTGTGATGACCTGCTCCAAAGATGCTTCATGCTGTATGGATATGGCTCC 1197

Db 500 AATGGCATGTGTGATGACCTGCTCCAAAGATGCTTCATGCTGTATGGATATGGCTCC 559

QY 1198 CCAACTGACATTTACCCCTCCGAGGCTGCTGTCGAGACACC-GAGCTGCTGTCATGTTG 1256

Db 560 CCAACTGACATTTACCCCTCCGAGGCTGCTGTCGAGACACC-GAGCTGCTGTCATGTTG 619

QY 1257 AGACTTGATGACAGTACATTTGTTTCTGATCTGGGG--ATGACATATAAGGTATG 1313

Db 620 AGACTTGATGACAGTACATTTGTTTCTGATCTGGGGATGAAACATATAAGGTATG 679

QY 1314 GAACA-CAACTGATGATGATTTGTTT--AAGGACCTTAATGACACAA 1359

Db 680 GAACACCAAGTACTTGTGAATTTTGTAGAGACCTTAATGACACAA 729

RESULT 6  
BE296484 758 bp mRNA linear EST 20-JUL-2000  
LOCUS 601174756F1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:3530341 5',  
DEFINITION mRNA sequence.  
ACCESSION BE296484  
VERSION BE296484.1 GI:9180046  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 758)  
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: ARCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at: Image.lnl.gov  
Plate: LNCM202 row: 0 column: 14  
High quality sequence stop: 617.

## FEATURES

source  
1. 758  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3530341"  
/clone\_id="NIH\_MGC\_17"  
/tissue\_type="rhodomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: muscle; Vector: pORF7; Site: 1: EcoRI;  
Site: 2: XhoI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GCCACGAG(9). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."  
BASE COUNT 212 a 149 c 202 g 195 t  
ORIGIN

Query Match 28.5%; Score 613.2; DB 10; Length 758;  
Best Local Similarity 96.3%; Pred. No. 1.4e-112;  
Matches 649; Conservative 0; Mismatches 23; Indels 2; Gaps 2;

QY 602 AATGTACGAGTACCTCTGATGACATGCTGTGAGAGAGCTTATGAGAGATGTCA 661  
Db 1 AATGTACGAGTACCTCTGATGACATGCTGTGAGAGAGCTTATGAGAGATGTCA 60  
QY 662 GGACAGATTCTCTGTGGAGAGCCCTGGCAGAGAGAGATGGGACAGATTATTCA 721

Db 61 GGACAGATTCTCTGTGGAGAGCCCTGGCAGAGAGAGATGGGACAGATTATTCA 120

QY 722 AAAACAAACCTCTGACGGGAATGCTCCCAACCTCTTTTATAGACACTTTATCTTA 781

Db 121 AAAACAAACCTCTGACGGGAATGCTCCCAACCTCTTTTATAGACACTTTATCTTA 180

QY 782 AAATATACAGACATTTGAGACATAGATGATTAATGAGATGTCGAGACATAGTTAC 841

Db 181 AAATATACAGACATTTGAGACATAGATGATTAATGAGATGTCGAGACATAGTTAC 240

QY 842 AGAGATTCTGCTCCGAGTGAACAGAGAGTCTTACTGTTTACAGATATGATGATC 901

Db 241 AGAGATTCTGCTCCGAGTGAACAGAGAGTCTTACTGTTTACAGATATGATGATGATC 300

QY 902 AGAAATAGTAAAGGCGCTTCAGACACACATCAATCAATCTGGATTAACAAACATTGG 961

Db 301 AGAAATAGTAAAGGCGCTTCAGACACACATCAATCAATCTGGATTAACAAACATTGG 360

QY 962 AATGCAAGGGAATTTCTACAGAGCCATACAGTTCAAGTCTGCTCCAGATATGAGA 1021

Db 361 AATGCAAGGGAATTTCTACAGAGCCATACAGTTCAAGTCTGCTCCAGATATGAGA 420

QY 1022 GAGTGTCAATACAGATATGAGATTCACAGCTCAGAGTGTGGATGT -AATACAGGT 1080

Db 421 GAGTGTCAATACAGATATGAGATTCACAGCTCAGAGTGTGGATGTAAATACAGGT 480

QY 1081 GAAATGCTAAACAGCTTGATTCACATTTGTAAGAGTCTGCTGCTTCAATAT 1140

Db 481 GAAATGCTAAACAGCTTGATTCACATTTGTAAGAGTCTGCTGCTTCAATAT 540

QY 1141 GGCATGATGATGACCTGCTCCAAAGATGCTTCCATTTGCTATGAGATATGCTCCCA 1200

Db 541 GGCATGATGATGACCTGCTCCAAAGATGCTTCCATTTGCTATGAGATATGCTCCCA 600

QY 1201 ACTGACATTAACCTCCGAGGAGGCTGCTGTC -GGACACGAGCTGCTGATGTTGAGA 1259

Db 601 CTGGACATTAACCTCCGAGGAGGCTGCTGTCGGACACAGAGCTGCTGATGTTGAGA 660

QY 1260 CTTTGTATGACNA 1273

Db 661 CTTTGTATGACNA 674

RESULT 7  
BE269297 692 bp mRNA linear EST 13-JUL-2000  
LOCUS 601186356F1 NIH\_MGC\_8 Homo sapiens cDNA clone IMAGE:3544549 5',  
DEFINITION mRNA sequence.  
ACCESSION BE269297  
VERSION BE269297.1 GI:9142915  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 692)  
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at: Image.lnl.gov  
Plate: LNCM239 row: 0 column: 14  
High quality sequence stop: 610.

## FEATURES

source  
1. 692  
Location/Qualifiers  
/organism="Homo sapiens"



LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT																														
Db 548	TATCTTAAACCATTTGTGCAGAGAGATTTTCATTAACGTCTCTGCAGGCTGGGGATTGGAT	532	CATATCGCTGAGAACATTTCTGTCACTAACCCTGGATGCGCAAAATCACTATGTCTCTGAACCT	591	Db 608	CATATTTGTTGGAACAAATTTCTGTCAATACCTGGATGCGCAAAATCACTATGTCTCTGAACCT	667	592	GTGTCGAAGGAATGCTTACCGAGTACCTCTGTATGGCATGTCTGTGAAGAAGCTTATCGAG	651	668	GTGTCCAAAGGAATGTTACCGAGTACCTCTGTATGGCATGTCTGTGAAGAAGCTTATCGAG	727	652	AGAAATGTCACGACAGATTTCTGTGTGAGAGAGCCCTGGCGAAGCAAGAGAGA--TGGGGA	708	728	AGAAATGTCACGACAGATTTCTGTGTGAGAGAGCCCTGGCGAAGCAAGAGAGATGCGGAC	787	709	CAGTATTTATTCAAAACCAACCTCTCTGACGGGAATG	745	788	CAGTATTTATTCAAAACCTCTCTGACGGGAATG	822																
RESULT 9	BO948186	1147 bp	mRNA	linear	EST	21-AUG-2002	AGENCOURT_8878641 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6464639	5', mRNA sequence.	BO948186	EST.	BO948186.1	GI:22363664	human.	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	1 (bases 1 to 1147)	NIH-MGC http://mgi.nci.nih.gov/.	National Institutes of Health. Mammalian Gene Collection (MGC) Unpublished (1999)	Contact: Robert Strausberg, Ph.D. Email: gsabbs-remail.nih.gov	Tissue Procurement: ATCC	cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNI). DNA sequencing by: Agencourt Bioscience Corporation	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNI at: http://image.llni.gov	Plate: L14M13986	row: a	column: 24	High quality sequence stop: 501.	location/Qualifiers	1. 1147	/organism="Homo sapiens"	/db_xref="taxon:9606"	/clone="IMAGE:6464639"	/clone_lib="NIH_MGC_71"	/tissue_type="leiomyosarcoma"	/lab_host="DH10B (phage-resistant)"	/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."	BASE COUNT	324 a	289 c	284 g	250 t	ORIGIN
Query Match	27.3%	Score 586.6;	DB 14;	Length 1147;	Best Local Similarity	97.6%;	Prod. No. 1.7e-164;	Matches 606;	Conservative	0;	Mismatches 14;	Indels 1;	Gaps 1;																												
Db 150	TTTCAGAAATTCCTCGAGAGAGAAACCTGTATTAATGCGAAACCCCTTAGGAAGATAA	209	172	CCAGAGAAGAAATTCCTAGACAGACATACAACACCTGTGCGACAGCTGTCTTAACCA	231	210	CCAGAGAAGAAATTCCTAGACAGACATACAACACCTGTGCGACAGCTGTCTTAACCA	269	232	GAACAGATATGTTTACGACAGACCTGCTATGAAGACTGAGAAATTTGTGGCCAAAACAAA	291																														

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Db	270	GAAGACGATATGTTTGGCAAGCAGCTGCTATGAAGAGCTGAGAAATTTGTGGCCAAAAA	329									
QY	292	CTTGGCAATGGACATCTTCCAGTATGATTTGGGCCAAGCAACGAAACTCTCAGCAAGCTAT	351									
Db	330	CTTGGCAATGGACATCTTCCAGTATGATTTGGGCCAAGCAACGAAACTCTCAGCAAGCTAT	389									
QY	352	CAAAAAGAAAAGAACTGTGTCTCAAAATTCCTTTGAGCAGTGGTTCAGAGTCAGATCAAGTG	411									
Db	390	GAAAAAGAAAAGAACTGTGTCTCAAAATTCCTTTGAGCAGTGGTTCAGAGTCAGATCAAGTG	449									
QY	412	GAATTTGTGGCAACATCTTATATATCCAAATGTGTCTATTACCAACATGGGCACATTAACCTCG	471									
Db	450	GAATTTGTGGCAACATCTTATATCCAAATGTGTCTATTACCAACATGGGCACATTAACCTCG	509									
QY	472	TATCTTAAACCTATGTGGAGAGACATTTGCATTAACCTGCTCTCCAGCTCGGGGATTTGGAT	531									
Db	510	TATCTTAAACCTATGTGGAGAGATTTGCATTAACCTGCTCTCCAGCTCGGGGATTTGGAT	569									
QY	532	CATATCGCTGAGAACATCTCTGCATACACCTGGATGGCAATCACTATGTGCTCTGAATTT	591									
Db	570	CATATCGCTGAGAACATCTCTGCATACACCTGGATGGCAATCACTATGTGCTCTGAATTT	629									
QY	592	GTGTGCAGAGAAATGTTACCGAGTGCACCTCTGATGGCATGCTGT -GGAAGACCTTATTCGA	650									
Db	630	GTGTGCAGAGAAATGTTACCGAGTGCACCTCTGATGGCATGCTGTGGGAAGAACCTTATTCGA	689									
QY	651	GAGATGTGTCAGACAGATCTCTGTGGAGAGAGCTGGGCAAGCAAGAGATGGGGACA	710									
Db	690	GAGATGTGTCAGACAGATCTCTGTGGAGAGAGAGCTGGGCAAGCAAGAGATGGGGACA	749									
QY	711	GTATTTATTCAAAAACAAC 731										
Db	750	GTATTTATTCAAAAACAAC 770										
RESUT 10	86829037	B6829037 743 bp mRNA linear EST 22-MAY-2001										
LOCUS	602753364F1	NIH_MGC_17 Homo sapiens CDNA clone IMAGE:4906215 5'										
DEFINITION	B6829037	mRNA sequence.										
ACCESSION	B6829037											
VERSION	B6829037.1	GI:14176624										
KEYWORDS	EST.											
SOURCE	human.											
ORGANISM	Homo sapiens											
REFERENCE	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;											
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.											
TITLE	1 (bases 1 to 743)											
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/											
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)											
	Unpublished (1999)											
	Contact: Robert Strausberg, Ph.D.											
	Email: cga@bbs.femail.nih.gov											
	Tissue Procurement: ATCC											
	CDNA Library Preparation: Ling Hong/Rubin Laboratory											
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)											
	DNA Sequencing by: Incyte Genomics, Inc.											
	Clone distribution: MGC clone distribution information can be											
	found through the I.M.A.G.E. Consortium/LNL at:											
	http://image.llnl.gov											
	Plate: LICM1808 row: 0 column: 16											



Db	641	GTGCAAGAAATGTCACCGAGTACCTTCGATGGCATGCTGTGGAAGAAACCTTATCGAGAG	700
Oy	654	AATGCTCAGGACAGATTTCTCTGTGGAGAGCGCTGGCAGACAGCAAGAGG	701
Db	701	AATGGCTCAGGACAGATTTCTCTGTGGAGAGCGCTGGCAGACAGCAAGAGG	748
RESULT 12			
LOCUS	BS518761	789 bp	linear
DEFINITION	602578346f1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3491843 5'		
ACCESSION	BS518761		
VERSION	BS518761.1	GI:13513945	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 789)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cga@bts-remail.nih.gov		
	Tissue Procurement: Gilbert Smith, Ph.D.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: LLM8536 row: k column: 12		
	High quality sequence stop: 587.		
FEATURES			
SOURCE	location/Qualifiers		
	1..789		
	/organism="Mus musculus"		
	/strain="CZECH 11"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:3491843"		
	/clone.lib="NCI CGAP_Lu29"		
	/tissue.type="spontaneous tumor, metastatic to mammary.		
	stem cell origin."		
	/lab.host="DH10B"		
	/note="Organ: Lung; Vector: pCMV-sPORT6; Site_1: SalI;		
	Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.		
	Library constructed by Life Technologies. Investigator		
	providing samples: Gilbert Smith, NIH"		
BASE COUNT	257 a 176 c 202 g 154 t		
ORIGIN			
Query Match	24.8%; Score 534.4; DB 12; Length 789;		
Best Local Similarity	86.1%; Pred. No. 7e-149;		
Matches 661; Conservative	0; Mismatches 101; Indels 6; Gaps 6		
Oy	65	CGATATGAGACCCGCCGAGGCGGTGCTCCAGAGAGAGCAAGCAAGTTAAGATTCCT	124
Db	1	CGATATGAGACCCGCCGAGGCGGTGCTCCAGAGAGAGCAAGCGCTTAAGTATGAATTCCT	60
Oy	125	CAGAGAGAGAAAGCTGTATTAATGCGGAACCCCTTGAGAAAGATTAATACAGAGAAAT	184
Db	61	CAGAGAGAGAAAGCTGTATTAATGCGGAACCCCTTGAGAAAGATTAATACAGAGAAAT	120
Oy	185	CACCTTAGACAGACATACAAACAGCTGTGCCAGACCTCTGCTTAACCAAGAAACAGTATT	244
Db	121	CACCTTAGACAGACATACAAACAGCTGTGCCAGACCTTTGCAATAACCAAGAGACAGTAT	180
Oy	245	TAGCAAGCACTGCTATGAAGACTGAGAAATTTGTGTGGCCAAACAAACTTGCATGCGA	304
Db	181	TAAACAAGCACTGCTATGAAGACTGAGAAATTTGTGTGGCCAAACAAACTTGCATGCGA	240
Oy	305	CTTCCAGATGATTTGTGCCCAAGCAAGCAAACTCTCAGCAAGCTATGAAAAGGAA- AAG	363

Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Db	241	CTTCCAGCATGATGTTGGCCCAAGCAGCGGAAACCTCTCAGCAAGCTATGAGAAACGA										
Qy	364	GAACGTGTGTGTCMAATACTTTGAGCAGTGGTGCAGAGTCAAGTGG -AATTGTGTA										
Db	301	GAGCGTGTGTCAACTATTTTGGAGCAGTGGTGCAGTGTGATCAAGTGAATTTGTGTA										
Qy	423	ACATCTTATATCCCAATAGTGTGTCATTTACCACATGGCAGCATAACTGTAATCTTAAC										
Db	361	ACACCTTATATCCCAATAGTGTGTCATTTACCACATGGCAGCATAACTGTAATCTTAAC										
Qy	483	TATGTTGCAAGAGATTTTCATPACTACTGCTGTCCACCTCGGGATGTGATCATATCGTGA										
Db	421	TATGCTGCAGAGGGGTTTCATTAACGTCACTGCTCCACACAGGGGGTGTGGACCACTGCTGA										
Qy	543	GAACCTTGTGCATATCCGATGCGAATGCCAATTCATATGCTGTGAACCTTGTGTGCAAGA										
Db	481	GAACCTTGTGCATATCCGATGCGAATGCCAATTCATATGCTGTGAACCTTGTGTGCAAGA										
Qy	603	ATGCTACCGC -AGTGACCTCTGTATGGCATGCTGTGGAAGACCTTATTCAGAGAAATGTCA										
Db	541	ATGCTACCGCAGTGTACGTGCGAGCGCATGTGTGAAAAGGCTCATGAGAGATGTGCA										
Qy	662	GGAC -AGATTTCTCTGTGA -GAGCGCTGCGACAGCAAGAGAGATGGGACAGTATTTT										
Db	601	GGACGGCTACTCTGTGGACGAGCGCCGACAGCGCAGAGGCTGGGACAGACTTACTTC										
Qy	720	CAAAACCAACCTCTGACGGGATAGTCCTCCCAACTCTTTTATAGACACTTATCC										
Db	661	AAAACCAAACTCATGTATGAGAACAGCGCTTCGCCAAAGACCAATTAAGACGCGTTAAC										
Qy	780	TAAATTTATACAGACATTGAGACATATGAATCTAATTTGAGAGATGG										
Db	721	TAAATTCATTA -AAGACATAGAGACATTAAGGTCATGAGATGGG										
RESULT 13												
LOCUS	BI103092	637 bp	mRNA	linear	EST 26 -JUN-2001							
DEFINITION	60288932JF1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5044684											
ACCESSION	BI103092											
VERSION	BI103092.1	GI:14553985										
KEYWORDS	EST.											
SOURCE	house mouse.											
ORGANISM	Mus musculus											
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;											
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.											
TITLE	1 (bases 1 to 637)											
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.											
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)											
	Unpublished (1999)											
	Contact: Robert Strausberg, Ph.D.											
	Email: cga@bbs.femail.nih.gov											
	Tissue Procurement: Jeffrey E. Green, M.D.											
	cDNA Library Preparation: Life Technologies, Inc.											
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)											
	DNA Sequencing by: Incyte Genomics, Inc.											
	Clone distribution: MGC clone distribution information can be											
	found through the I.M.A.G.E. Consortium/LML at:											
	http://image.llnl.gov											
	Plate: L14M1122 row: 1 column: 05											

Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI\_CGAP Library. 1"

BASE COUNT 185 a 148 c 172 g 132 t  
 Query Match 24.8%; Score 532.4; DB 13; Length 637;  
 Best Local Similarity 90.3%; Pred. No. 2.5e-148;  
 Matches 569; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Am Klopferspitz 18a D-82152 Martinsried, Germany  
 This is the 3' sequence of the clone insert  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 Braunschweig/Germany) within the cDNA sequencing consortium of the  
 German Genome Project.  
 This clone (DKFZ434D2317) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
 Location/Qualifiers

## FEATURES

Source

1. 544  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="DKFZ434D2317"  
 /clone\_id="434 (synonym: htes3)"  
 /tissue="testis"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Vector: pSport1; Site\_1: NotI; site\_2: SalI"

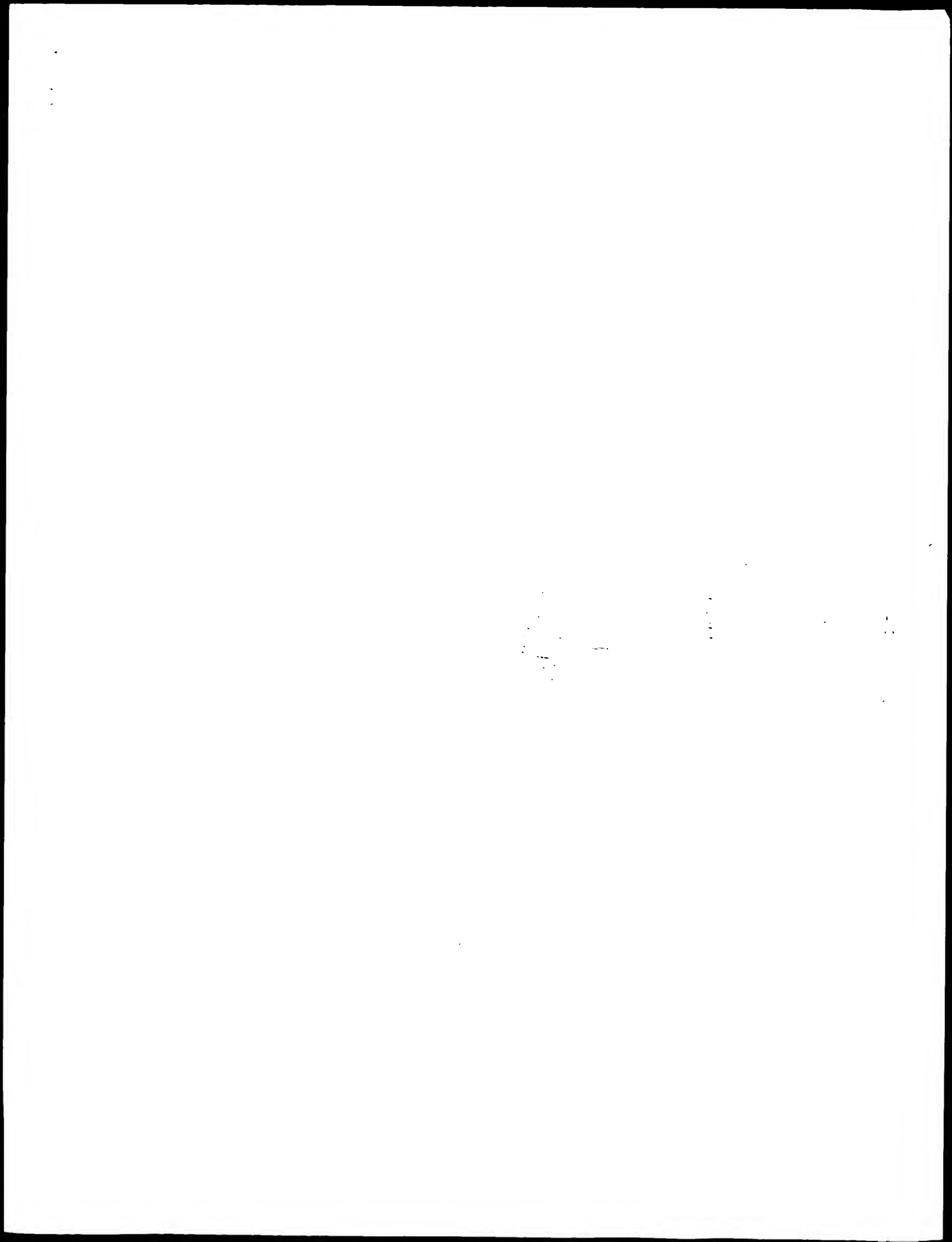
BASE COUNT 151 a 102 c 153 g 138 t  
 Query Match 24.6%; Score 529.8; DB 9; Length 544;  
 Best Local Similarity 99.4%; Pred. No. 1.4e-147;  
 Matches 542; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 26 AGGCGCGCCCGCGGAGAGACCGGACAGTGGCCCTCGGCGGATTTGAGCCCGCGAGG 85  
 DB 4 ACCGGTGGCCCGCGGAGAGACCGGACAGTGGCCCTCGGCGGATTTGAGCCCGCGAGG 63  
 QY 86 CGGTCTCCGAG 145  
 DB 64 CGGTCTCCGAG 123  
 QY 146 ATGGCGAACCCTCAGAGAGATTAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 205  
 DB 124 ATGGCGAACCCTCAGAGAGATTAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 183  
 QY 206 GCTGTGCCAGACTCTGCTTAACCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 265  
 DB 184 GCTGTGCCAGACTCTGCTTAACCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 243  
 QY 266 CTGAGATTTGTTGGCCAAACAAACAACTTGGCAATGGCACTTGCAGATGATTTGCC 325  
 DB 244 CTGAGATTTGTTGGCCAAACAAACAACTTGGCACTTGCAGATGATTTGCC 303  
 QY 326 AGCAACGAAACTCTCAGCAAGATGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 385  
 DB 304 AGCAACGAAACTCTCAGCAAGATGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 363  
 QY 386 AGCAGTGGTCAAGATCAAGATGAAATTTGGCAACCTTTATATCCCAATGTC 445  
 DB 364 AGCAGTGGTCAAGATCAAGATGAAATTTGGCAACCTTTATATCCCAATGTC 423  
 QY 446 ATTACCAACATGGGACATTAACCTGATTTAACTTATGTCAGAGAGATTTTCATTA 505  
 DB 424 ATTACCAACATGGGACATTAACCTGATTTAACTTATGTCAGAGAGATTTTCATTA 483  
 QY 506 CTGCTCTGCCAGCTCGGGGATTTGGATCATATCCCTGGAACATTTCTGCTGAGT 565  
 DB 484 CTGCTCTGCCAGCTCGGGGATTTGGATCATATCCCTGGAACATTTCTGCTGAGT 543  
 QY 566 CCAAAATCACTATGCTGCTGTAAGCTTGTGCAAGAAAGTGTACCGAGTACCTGTG 625  
 DB 544 CCAAAATCACTATGCTGCTGTAAGCTTGTGCAAGAAAGTGTACCGAGTACCTGTG 603  
 QY 626 GCATGCTGTGGAAGAAAGCTTATCGAGAGAA 655  
 DB 604 GCATGCTGTGGAAGAAAGCTTATCGAGAGAA 633

RESULT 14  
 LOCUS AL041438/c 544 bp mRNA linear EST 29-FEB-2000  
 DEFINITION DKFZ434D2317\_s1 434 (synonym: htes3) Homo sapiens cDNA clone  
 ACCESSION AL041438  
 VERSION AL041438.1 GI:5420789  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 544)  
 Bloeker, H., Boecker, M., Brandt, P., Mewes, H. W., Gassenhuber, J., and  
 Wiemann, S.  
 EST (Bloeker, et al.)  
 Unpublished (1999)  
 COMMENT CONTACT: Bloeker H  
 MIPS

RESULT 15  
 LOCUS B145545 888 bp mRNA linear EST 21-AUG-2001  
 DEFINITION 603173974F1 NCI\_CGAP\_Mam5 Mus musculus cDNA clone IMAGE:5253090 5',







|||||  
Db 241 TGTTCAGAACCTGCTATGAGACATGAGATGTGTGGCCAAACAACTTGCCAT 300  
QY 301 GGCACTTCAGATGATGTTGGCCAGCAAGGAACTGTCAGAACCTTATGAAAGAA 360  
Db 301 GGCACTTCAGATGATGTTGGCCAGCAAGGAACTGTCAGAACCTTATGAAAGAA 360  
QY 361 AAGGAACTGTGTGTCAAATCTTGTAGCAGTGTGTGAGAGTCAATCAATGATTTGTG 420  
Db 361 AAGGAACTGTGTGTCAAATCTTGTAGCAGTGTGTGAGAGTCAATCAATGATTTGTG 420  
QY 421 GAACATCTTATATCCCAATATGTGATATCCAAATGAGCAATATCTGTATCTTAA 480  
Db 421 GAACATCTTATATCCCAATATGTGATATCCAAATGAGCAATATCTGTATCTTAA 480  
QY 481 CCTATGTTGAGAGAGATTCATTAATGCTGTGTGACCTGCGGGGATTTGATATGCT 540  
Db 481 CCTATGTTGAGAGAGATTCATTAATGCTGTGTGACCTGCGGGGATTTGATATGCT 540  
QY 541 GAGAACATTTGTGATACCTGATGCTGCAATCACTATGCTGCTGCAATCTTGTGCAAG 600  
Db 541 GAGAACATTTGTGATACCTGATGCTGCAATCACTATGCTGCTGCAATCTTGTGCAAG 600  
QY 601 GAATGTGACGAGAGACCTGCTGATGCTGCTGTGAGAGAGCTTATCGAGAGATGCTC 660  
Db 601 GAATGTGACGAGAGACCTGCTGATGCTGCTGTGAGAGAGCTTATCGAGAGATGCTC 660  
QY 661 AGGACAGATTTCTGTGAGAGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
Db 661 AGGACAGATTTCTGTGAGAGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
QY 721 AAAAACAACTCTCTGAG 780  
Db 721 AAAAACAACTCTCTGAG 780  
QY 781 AAAATTAACAAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840  
Db 781 AAAATTAACAAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840  
QY 841 CAGAGATTCCTGCTGAG 900  
Db 841 CAGAGATTCCTGCTGAG 900  
QY 901 CAGAGATTCCTGCTGAG 960  
Db 901 CAGAGATTCCTGCTGAG 960  
QY 961 GAATGTGAG 1020  
Db 961 GAATGTGAG 1020  
QY 1021 AGAGTATCATTAACAGAGATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTC 1080  
Db 1021 AGAGTATCATTAACAGAGATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTC 1080  
QY 1081 GAAATCTTAAACAGTGTGATTCACATTTGTGAAGCAGTTCTGCTGCTTCAATAT 1140  
Db 1081 GAAATCTTAAACAGTGTGATTCACATTTGTGAAGCAGTTCTGCTGCTTCAATAT 1140  
QY 1141 GGCATGATGTCAGCTGCTGCAAAAGATGCTTCAATTTGTGAGATTTGTGAGAG 1200  
Db 1141 GGCATGATGTCAGCTGCTGCAAAAGATGCTTCAATTTGTGAGATTTGTGAGAG 1200  
QY 1201 ACTGACATTCACCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
Db 1201 ACTGACATTCACCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
QY 1261 TTTGATGACAGATTCATTTGCTGATCTGAGAGATGAGAGATTAAGATTTGAGAG 1320  
Db 1261 TTTGATGACAGATTCATTTGCTGATCTGAGAGATGAGAGATTAAGATTTGAGAG 1320  
QY 1321 AGTACTTGTGATTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
Db 1321 AGTACTTGTGATTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380

Db 1321 AGTACTTGTGATTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
QY 1381 TACAGGAG 1440  
Db 1381 TACAGGAG 1440  
QY 1441 GAATGTGTCAGATTTGTCAGATTTGTCAGATTTGTCAGATTTGTCAGATTTGTC 1500  
Db 1441 GAATGTGTCAGATTTGTCAGATTTGTCAGATTTGTCAGATTTGTCAGATTTGTC 1500  
QY 1501 TTTGATTAACAGAGATTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
Db 1501 TTTGATTAACAGAGATTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
QY 1561 GTGGTGTCTTTGAG 1620  
Db 1561 GTGGTGTCTTTGAG 1620  
QY 1621 CATTCGGAAGAGATTTTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 1680  
Db 1621 CATTCGGAAGAGATTTTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 1680  
QY 1681 GATGACAGAGATTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
Db 1681 GATGACAGAGATTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
QY 1741 GCTTCCCTCTCTGAG 1800  
Db 1741 GCTTCCCTCTCTGAG 1800  
QY 1801 ACTTCCGAG 1860  
Db 1801 ACTTCCGAG 1860  
QY 1861 AACACAGATTAACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1920  
Db 1861 AACACAGATTAACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1920  
QY 1921 TGAGAGCTCTTTGAG 1980  
Db 1921 TGAGAGCTCTTTGAG 1980  
QY 1981 ACTGACTCTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2040  
Db 1981 ACTGACTCTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2040  
QY 2041 TTTTAAACCTTCCCT 2100  
Db 2041 TTTTAAACCTTCCCT 2100  
QY 2101 CAGACAAAGAGAGATTTAAATATATTTAGCTTTTCCAGAGAGAGAGAGAGAGAG 2151  
Db 2101 CAGACAAAGAGAGATTTAAATATATTTAGCTTTTCCAGAGAGAGAGAGAGAGAG 2151

RESULT 2  
US-10-023-530-1  
: Sequence 1, Application US/10023530  
: Publication No. US2003007956A1  
: GENERAL INFORMATION:  
: APPLICANT: LEGRAND, Pierre  
: APPLICANT: BERNARDUS, Richard  
: APPLICANT: BIOT, Guillaume  
: APPLICANT: LASSOT, Irina  
: TITLE OF INVENTION: PROTEINS THAT INTERACT WITH BETA TRCP  
: FILE REFERENCE: B4717A  
: CURRENT APPLICATION NUMBER: US/10/023, 530  
: CURRENT FILING DATE: 2002-04-22  
: PRIOR APPLICATION NUMBER: 60/256, 276  
: PRIOR FILING DATE: 2000-12-18  
: NUMBER OF SEQ ID NOS: 30  
: SOFTWARE: PatentIn version 3.1  
: SEQ ID NO 1











Query Match	5.8%	Score 124.2;	DB 9;	Length 2010;
Best local Similarity	50.2%	Pred. No. 2.1e-27;		
Matches 340; Conservative	0;	Mismatches 328;	Indels 9;	Gaps 1

```

Yy 1543 ATTAAGTGTGGGATCT 1559
      ||||| ||||| |
Db 1567 GTTAAATCTGGGATAT 1583

RESULT 12
US-09-328-877A-24
: Sequence 24, Application US/09328877A
: Patent No. US20020177187A1
: GENERAL INFORMATION:
: APPLICANT: Gurney, Mark E.
: APPLICANT: Li, Jihne
: APPLICANT: Pauley, Adele M.
: APPLICANT: Pharmacia & Upjohn Company
: TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
: FILE REFERENCE: 6142
: CURRENT APPLICATION NUMBER: US/09/328,877A
: CURRENT FILING DATE: 1999-06-09
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 24
: LENGTH: 2010
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: VEHIS tagged
US-09-328-877A-24

Query Match      5 8%; Score 124.2; DB 9; Length 2010;
Best Local Similarity 50.2%; Pred. No. 2,1e-27;
Matches 340; Conservative 0; Mismatches 328; Indels 9; Gaps

Yy 883 TGTTCACATATGATGATCAGAAATAGTAAAGCGGCTTCGAGACAAACATCAAGATC 942
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 916 TCGTTCACATTTTGTGTACCGAATAGTTAGTGTTGTGATGACACACTTTAAAGTT 975

Yy 943 TGGGATAAAAACACATGGAATGCAAGCAATTTTCACAGGCCATACAGGTTCAAGTCTC 1002
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 976 TGGTCACGCTACAGGCAAAATGTCTGAGAACATTAGTGGGACATACAGGTGAGATATG 1035

Yy 1003 TGTCTCCATATGATGAGAGAGTGTATCAATACAGATCATCGATTCCACGGTCAAGAG 1062
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1036 TCATCCAAATGAGAGACAAACATCATCATATGATGATCTACAGATCGGACACTCAAGTG 1095

Yy 1063 TGGGATGTAAATACAGGTGCAAAATGCTAAACACGTTGATTCACCATTGTGAACGACTT 1122
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1096 TGGAAATGACAGACGTGGAAATGTATACACACTTATATGGCATCTCCACTGTGCT 1155

Yy 1123 CACTTGCCTTCAATAAATCCGATGATGGTGAGCTGCTCCAAAGATCGTTCCATGCTGTA 1182
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1156 TGTATGCATCTTCATGAAAAAAGATTGTTAGCGGTTCTCGAGATGCCACTCTTAGGGTT 1215

Yy 1183 TGGGATATGGCTCCCAACTGACATTAACCTCGGAGGGTGTGTGGAGACCCGAGCT 1242
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1216 TGGGATATGAGACGAGCGCAGTGT-----ACATGTTTATGGGTTCATGTTGCA 1266

Yy 1243 GCTGTCAATGTTGTAGACTTCATACAGATGATGATGTTGTCGATTCGGGATAGACT 1302
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1267 GCAGTCCGCTGTGTTCAATATGATGAGGAGGAGGTTGTATGAGACATATGATTTTATG 1326

Yy 1303 ATTAAGGATGACACACAGTACTTGTGAATTTGTGAAGACCTTAAATGACACAAACGA 1362
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1327 GTAAAGGATGTGGGATCCAGAGACTAAACCTGTCTACACAGCTTGACAGGGCATACAT 1386

Yy 1363 GGCATTTGCTGTTTGCATATACAGGACAGGCTGTGATGATGCTCATCTGACAAACT 1422
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1387 AGAGTCTTATTAATACATTTGATGATGATTCATGATGGAGATGATCTCTTGAATACATCA 1446

Yy 1423 ATCAGATTTGAGGACATGGAATGTGCTGATGTTTACAGATGTTGAAGGCCATGAGAA 1482
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



OY	1243	GCTGCATGCTTTGGAGACTTGTATGACAAATACATCTGTTTCGCAATCTGGGATAGACT	1302
Db	1311	GCAGTCCCTCTGTTCCATATGATGATGGCAGAGSGTTCTTAGTGGACATATGATTTTATG	1370
OY	1303	ATTAAGGTATGGAAACACACAGTACTTGTGAATTTGTAAAGACCTTAATGACACAAACGA	1362
Db	1371	GTAAGAGGTGGTGATCCAGAGACTGAACAACCTGTCTACACAGCTTCAGGGGACTACTAAT	1430
OY	1363	GGCATTTGCTTTTGGACGTACAGGAGACAGCTGCTGTAGTGGCCATCTCACAAACT	1422
Db	1431	AGAGTCTATTCATTACAGTTTTCATGTGATGCCATGTGCTGATGATCTCTTGATTCATCA	1490
OY	1423	ATTCAGATTATGGGACATAGATGTGTCATGTTTACGAGTGTTCGAAAGGCCATGAGAA	1482
Db	1491	ATCCGTGTTTGGATGTGGAGACAGGGAATTGGATTCCACACCTTAAACAGGGACACAGTCG	1550
OY	1483	TTTGGCTGCTGTATATCCATTTCGATTAACACAGAGATAGTCACTGGGCCCATGTGAA	1542
Db	1551	TTTACAAAGTGAATGAGAACTCAAGAACATATTTCTGTCTCTGGGAATGCAATCTTACA	1610
OY	1543	ATTAAGCTGGGACTT	1559
Db	1611	GTTAAAACTGGGATAT	1627

```

RESULT 15
US-09-213-888-2
: Sequence 2, Application US/09213888A
: Patent No. US20020164683A1
: GENERAL INFORMATION:
: APPLICANT: Gurney, Mark E.
: APPLICANT: Li, Jinhe
: APPLICANT: Pauley, Adele M.
: APPLICANT: Pharmacia & Upjohn Company
: TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
: FILE REFERENCE: 6142
: CURRENT APPLICATION NUMBER: US/09/213,888A
: CURRENT FILING DATE: 1998-12-17
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 3571
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (2506)
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (3393)
: US-09-213-888-2

```

Query Match	5.88;	Score 124,2;	DB 9;	Length 3571;
Best Local Similarity	50.28;	Pred. No. 3.1e-27;		
Matches 340;	Conservative	0;	Mismatches 328;	Indels 9;
				Gaps 1

QY	883	TGTTTACAGTATATATCACTCGAAATAGTAAACGGGCTTCGAGAACACCAATCAAGATC	942
Db	981	TGCTTACAGTTTGTGTGTAACCGAATATAGTTAGTGGTTCGTGATGACACACTTTAAAGTT	1040
QY	943	TGGGATTAACACACATTGGAGTAAGCAAGCAATTTCTCACGGCCATACAGTTCACTCTTC	1002
Db	1041	TGGTCACGACATACAGGCAAAATGTCTGGAACACTTAGTGGACATTAACAGTGGACATTAGG	1100
QY	1003	TGTCCTCAGTATGATGAGAGATGATCATTAACAGATCATCTGGAAATTCACGGTCAAGATG	1062
Db	1101	TCATACCAAAATGAGAGACCAACATCAATCATTAAGTGTCTTACAGATCGACACTCAAAATG	1160
QY	1063	TGGGATTAATACAGGTGAATATGCTAAACACAGTTGATTCACCACTTTGAAGCAACTCTTG	1122
Db	1161	TGGAAATGACAGACTGAGAAATATATACACACTTATATAGGCAATATCCACACTGTGGT	1220

OY	1123	CACGTGCGTTCCATTAATGGCATGAGTGACCTGCTCCAAAGATCGTTCATTGCCTGA	1182
Dd	1221	TGTATGCATCTTCATGAAAAAGATTGTAGCGGTTTCGAGATGCCACTCTTTAGGGTT	1280
OY	1183	TGGCATATGGCCTCCGCCACTGCATTAACCCTCGGAGGGTCTGTGTGGACACCGAGCT	1242
Dd	1281	TGGGATATTGAGACACGGCCAGTGTTTT-----ACATGTTTTGATGGGTCAATTTTCCA	1331
OY	1243	GCTGTCAATGTTTGNAGCTTTGATGACAGTACATTGTTTCTTGCACTGGGGATAGAAT	1302
Dd	1332	GCAGTCGCGCTGTGTCAAATGATGGCAGAGGGGTGTAGTAGGAGCATATGATTTTATG	1391
OY	1303	ATAAAGATAGGAACAACAATACTGTGAAATTTGTAAAGACCTTTAAATGACACAAAAGA	1362
Dd	1392	GTAAGAAGTGGGATCCAGAGACTGAAACCTGTCACACACTTCGACGGGCATATCTAAT	1451
OY	1363	GGCATTGCCCTGTTTGCAGTACAGSGGACGGCTGTGATGAGTGCCATCTGCACAACT	1422
Dd	1452	AGAGTCTATTCATCACAGTTTGATGTATCCATGTGATGAGGAGATCTGTGATACATCA	1511
OY	1423	ATCAGATTAATGGGACATAGAAATGTGTGCATGTTTACGAGTGTGAAGGCCATGAGGAA	1482
Dd	1512	ATCCGTGTTTGGGATGTGGAGACAGGAAATTGCATTCCACAGCTTAAACAGGGGACACAGTCG	1571
OY	1483	TTTGCTGGCTGTATTCGATTGTATAACAGAGAGATAGTCACTGAGGGCCCTATAGTGAANA	1542
Dd	1572	TTAACAAAGTGAATGGAATCAACAACATATTCCTGTGCTGGGAATGCAAGATTCTTACA	1631
OY	1543	ATTAAGCTGGGATCT 1559	
Dd	1632	GTTAAATCTGGGATTA 1648	

Search completed: February 22, 2003, 02:00:32  
Job time : 185 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 09:55:27 ; Search time 17 Seconds  
(without alignments)  
984.803 Million cell updates/sec

Title: US-09-601-168b-2  
Perfect score: 3034  
Sequence: 1 MDPAAVLQEKALKEFMSSE.....PAAQAEPPSPSRITYTISR 569

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgcn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgcn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgcn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgcn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgcn2\_6/ptodata/1/1aa/PCRTUS.COMB.pep:\*  
6: /cgcn2\_6/ptodata/1/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2582.5	85.1	517	1	US-08-190-802A-30 Sequence 30, Appl
2	2582.5	85.1	517	4	US-08-477-346-30 Sequence 30, Appl
3	2582.5	85.1	517	4	US-08-473-089-30 Sequence 30, Appl
4	2582.5	85.1	517	4	US-08-487-072A-30 Sequence 30, Appl
5	545	18.0	640	4	US-09-177-165A-30 Sequence 30, Appl
6	520	17.1	587	3	US-08-899-578-2 Sequence 2, Appl
7	399	13.2	779	1	US-08-190-802A-32 Sequence 32, Appl
8	399	13.2	779	4	US-08-477-346-32 Sequence 32, Appl
9	399	13.2	779	4	US-08-473-089-32 Sequence 32, Appl
10	399	13.2	779	4	US-08-487-072A-32 Sequence 32, Appl
11	399	13.2	779	4	US-09-177-165A-29 Sequence 29, Appl
12	392	12.9	732	4	US-08-914-999-8 Sequence 8, Appl
13	354	11.7	409	2	US-08-283-917-3 Sequence 3, Appl
14	354	11.7	409	2	US-08-961-716-3 Sequence 3, Appl
15	354	11.7	410	2	US-08-283-917-9 Sequence 9, Appl
16	354	11.7	410	2	US-08-961-716-9 Sequence 9, Appl
17	339.5	11.2	409	1	US-08-190-802A-51 Sequence 51, Appl
18	339.5	11.2	409	4	US-08-477-346-51 Sequence 51, Appl
19	339.5	11.2	409	4	US-08-473-089-51 Sequence 51, Appl
20	339.5	11.2	409	4	US-08-487-072A-51 Sequence 51, Appl
21	337	11.1	1146	4	US-08-914-999-6 Sequence 6, Appl
22	321.5	10.6	514	1	US-08-190-802A-66 Sequence 66, Appl
23	321.5	10.6	514	4	US-08-477-346-66 Sequence 66, Appl
24	321.5	10.6	514	4	US-08-473-089-66 Sequence 66, Appl
25	321.5	10.6	514	4	US-08-487-072A-66 Sequence 66, Appl
26	318	10.5	422	1	US-08-190-802A-52 Sequence 52, Appl
27	318	10.5	422	4	US-08-477-346-52 Sequence 52, Appl

28	318	10.5	422	4	US-08-473-089-52	Sequence 52, Appl
29	318	10.5	422	4	US-08-487-072A-52	Sequence 52, Appl
30	313.5	10.3	209	3	US-08-899-578-6	Sequence 6, Appl
31	306	10.1	704	1	US-08-188-582-5	Sequence 5, Appl
32	306	10.1	704	1	US-08-646-715-5	Sequence 5, Appl
33	305	10.1	704	1	US-08-190-802A-62	Sequence 62, Appl
34	305	10.1	704	2	US-08-308-818-3	Sequence 3, Appl
35	305	10.1	704	4	US-08-477-346-62	Sequence 62, Appl
36	305	10.1	704	4	US-08-473-089-62	Sequence 62, Appl
37	305	10.1	704	4	US-08-487-072A-62	Sequence 62, Appl
38	299.5	9.9	1194	4	US-09-092-508-2	Sequence 2, Appl
39	299.5	9.9	1194	4	US-09-435-115-2	Sequence 2, Appl
40	299.5	9.9	1194	4	US-09-089-023-26	Sequence 26, Appl
41	299.5	9.9	1194	4	US-09-098-310-2	Sequence 2, Appl
42	299.5	9.9	1205	4	US-09-092-508-16	Sequence 16, Appl
43	299.5	9.8	1205	4	US-09-435-115-16	Sequence 16, Appl
44	298.5	9.8	704	1	US-08-188-582-18	Sequence 18, Appl
45	298.5	9.8	704	1	US-08-646-715-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1  
US-08-190-802A-30  
Sequence 30, Application US/08190802A  
Patent No. 5519003  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: Dehlinger & Associates  
STREET: P.O. Box 60850  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-0850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,802A  
FILING DATE: 01-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 8600-0139  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0860  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO.: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-190-802A-30  
Query Match 85.1% Score 2582.5; DB 1; Length 517;  
Best Local Similarity 91.4%; Pred. No. 2.9e-266;  
Matches 487; Conservative 7; Mismatches 8; Indels 31; Gaps 2;  
18 SSREDCNNNGEPPRKIIPEKNSLRQIYNSCARLCLNDEIVCLASTAMTEKCVAKTKLAN 77

PRINTED

```

Db 13 ASEREDCNDEPPRKITEKNTLRO-----TKLAN 42
QY 78 GTSSMIVPRQKRLSASYEKEKELCVKYEOWSESDQVEFEVHLISQCHYOHGHTINSYLK 137
Db 43 GTSSMIVPRQKRLSANEKEKELCVKYEOWSESDQVEFEVHLISQCHYOHGHTINSYLK 102
QY 138 PMLORDEFTALPARGLDHAENILSYLDAKSLCAELVCKEYRYTSDGMLMKKLIERNV 197
Db 103 PMLORDEFTALPARGLDHAENILSYLDAKSLCAELVCKEYRYTSDGMLMKKLIERNV 162
QY 198 RTDSLIRGLAERRGQVLEFKNPPDGNAPNSFYRALYPKTIIDITISNMRCGRSL 257
Db 163 RTDSLIRGLAERRGQVLEFKNPPDGNAPNSFYRALYPKTIIDITISNMRCGRSL 222
QY 258 QRHCSEFSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECKRVLGHTGSVLCLOYDE 317
Db 223 QRHCSEFSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECKRVLGHTGSVLCLOYDE 282
QY 318 RVITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMAMVTCSDRSIAVMDASA 377
Db 283 RVITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMAMVTCSDRSIAVMDASA 341
QY 378 TDITLRVYVGHRAAVNVVDFDOKIYVSGDRITIKVWNTSTCEVFTLNGHKGRIACLO 437
Db 342 TDITLRVYVGHRAAVNVVDFDOKIYVSGDRITIKVWNTSTCEVFTLNGHKGRIACLO 401
QY 438 YRDLVYVSGSSDNTIRLMDIEGACLRVLEGEHELVCIRFDNKRIVSGAYDGKIKYWDL 497
Db 402 YRDLVYVSGSSDNTIRLMDIEGACLRVLEGEHELVCIRFDNKRIVSGAYDGKIKYWDL 461
QY 498 VALDPRAPAGTLCRLTIVESHGVRFLQDFEFOIVSSSHDDTILIMDFLNDP 550
Db 462 VALDPRAPAGTLCRLTIVESHGVRFLQDFEFOIVSSSHDDTILIMDFLNDP 514

```

RESULT 2  
US-08-477-346-30  
; Sequence 30, Application US/08477346  
; Patent No. 6262023  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,346  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,072  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 30:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 517 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13
US-08-477-346-30
Query Match 85.1%; Score 2582.5; DB 4; Length 517;
Best Local Similarity 91.4%; Pred. No. 2.9e-266;
Matches 487; Conservative 7; Mismatches 8; Indels 31; Gaps 2;
Db 13 ASEREDCNDEPPRKITEKNTLRO-----TKLAN 42
QY 18 SSEREDCNDEPPRKITEKNTLRO-----TKLAN 42
Db 13 ASEREDCNDEPPRKITEKNTLRO-----TKLAN 42
QY 78 GTSSMIVPRQKRLSASYEKEKELCVKYEOWSESDQVEFEVHLISQCHYOHGHTINSYLK 137
Db 43 GTSSMIVPRQKRLSANEKEKELCVKYEOWSESDQVEFEVHLISQCHYOHGHTINSYLK 102
QY 138 PMLORDEFTALPARGLDHAENILSYLDAKSLCAELVCKEYRYTSDGMLMKKLIERNV 197
Db 103 PMLORDEFTALPARGLDHAENILSYLDAKSLCAELVCKEYRYTSDGMLMKKLIERNV 162
QY 198 RTDSLIRGLAERRGQVLEFKNPPDGNAPNSFYRALYPKTIIDITISNMRCGRSL 257
Db 163 RTDSLIRGLAERRGQVLEFKNPPDGNAPNSFYRALYPKTIIDITISNMRCGRSL 222
QY 258 QRHCSEFSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECKRVLGHTGSVLCLOYDE 317
Db 223 QRHCSEFSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECKRVLGHTGSVLCLOYDE 282
QY 318 RVITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMAMVTCSDRSIAVMDASA 377
Db 283 RVITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMAMVTCSDRSIAVMDASA 341
QY 378 TDITLRVYVGHRAAVNVVDFDOKIYVSGDRITIKVWNTSTCEVFTLNGHKGRIACLO 437
Db 342 TDITLRVYVGHRAAVNVVDFDOKIYVSGDRITIKVWNTSTCEVFTLNGHKGRIACLO 401
QY 438 YRDLVYVSGSSDNTIRLMDIEGACLRVLEGEHELVCIRFDNKRIVSGAYDGKIKYWDL 497
Db 402 YRDLVYVSGSSDNTIRLMDIEGACLRVLEGEHELVCIRFDNKRIVSGAYDGKIKYWDL 461
QY 498 VALDPRAPAGTLCRLTIVESHGVRFLQDFEFOIVSSSHDDTILIMDFLNDP 550
Db 462 VALDPRAPAGTLCRLTIVESHGVRFLQDFEFOIVSSSHDDTILIMDFLNDP 514

```

RESULT 3  
US-08-473-089-30  
; Sequence 30, Application US/08473089  
; Patent No. 6342368  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,089  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MORASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.22  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-473-089-30

Query Match 85.1%; Score 2582.5; DB 4; Length 517;  
Best Local Similarity 91.4%; Pred. No. 2.9e-266;  
Matches 487; Conservative 7; Mismatches 8; Indels 31; Gaps 2;

QY 18 SSREDCNNGEPPRIKTIPEKNSLRQTYNSCARLCLNGETVCLASTAMKTEKNCVAKTKLAN 77  
DB 13 ASEREDCNDEPPRIKTIPEKNSLRQTYNSCARLCLNGETVCLASTAMKTEKNCVAKTKLAN 42  
QY 78 GTSSMIVPKOKRLSAYEKEKELCVYFFQWSESDQVEFEVHLISOMCHYGHINSYLK 137  
DB 43 GTSSMIVPKOKRLSAYEKEKELCVYFFQWSESDQVEFEVHLISOMCHYGHINSYLK 102  
QY 138 PMQORFTTALPARGLDHAENILSTYDAKSLCAAEVCKEYRYTSDGMLMKLIERVY 197  
DB 103 PMQORFTTALPARGLDHAENILSTYDAKSLCAAEVCKEYRYTSDGMLMKLIERVY 162  
QY 198 RTSLMRGLAERGGWGYLFKNKPPDGNAPNSFYALPKTIIDETTESNMRCGRHSL 257  
DB 163 RTSLMRGLAERGGWGYLFKNKPPDGNAPNSFYALPKTIIDETTESNMRCGRHSL 222  
QY 258 QRIHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECKRIILGHTGSVLCQYDE 317  
DB 223 QRIHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECKRIILGHTGSVLCQYDE 282  
QY 318 RVITIGSSDSTVAVWVNTGEMNLTLIHCEAVLHLRFNNGMMVYCSKDRSLAVMDMASP 377  
DB 283 RVITIG-SDSTVAVWVNTGEMNLTLIHCEAVLHLRFNNGMMVYCSKDRSLAVMDMASP 341  
QY 378 TDTLRRVAVGHRAAVNVDFDKYIVSASGDTIKVWNTSTCEFYRTLNHGRGIACIQ 437  
DB 342 TDTLRRVAVGHRAAVNVDFDKYIVSASGDTIKVWNTSTCEFYRTLNHGRGIACIQ 401  
QY 438 YRDLVAVSGSSDNTIRLMDIEGACLRVLEGHBEVLCIRFDNKRIVSGAYDGKIKYWDL 497  
DB 402 YRDLVAVSGSSDNTIRLMDIEGACLRVLEGHBEVLCIRFDNKRIVSGAYDGKIKYWDL 461  
QY 498 VVALDPPAPAGTCLRTLVHEHSGVFRLODFQIVSSSHDDTLLMDPLNDP 550  
DB 462 VVALDPPAPAGTCLRTLVHEHSGVFRLODFQIVSSSHDDTLLMDPLNDP 514

RESULT 4  
US-08-487-072A-30  
Sequence 30, Application US/08487072A  
Patent No. 6423684  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
APPLICANT: Ron. Dorit  
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,072A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MORASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 2550-0025.20  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13  
US-08-487-072A-30

Query Match 85.1%; Score 2582.5; DB 4; Length 517;  
Best Local Similarity 91.4%; Pred. No. 2.9e-266;  
Matches 487; Conservative 7; Mismatches 8; Indels 31; Gaps 2;

QY 18 SSREDCNNGEPPRIKTIPEKNSLRQTYNSCARLCLNGETVCLASTAMKTEKNCVAKTKLAN 77  
DB 13 ASEREDCNDEPPRIKTIPEKNSLRQTYNSCARLCLNGETVCLASTAMKTEKNCVAKTKLAN 42  
QY 78 GTSSMIVPKOKRLSAYEKEKELCVYFFQWSESDQVEFEVHLISOMCHYGHINSYLK 137  
DB 43 GTSSMIVPKOKRLSAYEKEKELCVYFFQWSESDQVEFEVHLISOMCHYGHINSYLK 102  
QY 138 PMQORFTTALPARGLDHAENILSTYDAKSLCAAEVCKEYRYTSDGMLMKLIERVY 197  
DB 103 PMQORFTTALPARGLDHAENILSTYDAKSLCAAEVCKEYRYTSDGMLMKLIERVY 162  
QY 198 RTSLMRGLAERGGWGYLFKNKPPDGNAPNSFYALPKTIIDETTESNMRCGRHSL 257  
DB 163 RTSLMRGLAERGGWGYLFKNKPPDGNAPNSFYALPKTIIDETTESNMRCGRHSL 222  
QY 258 QRIHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECKRIILGHTGSVLCQYDE 317  
DB 223 QRIHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECKRIILGHTGSVLCQYDE 282  
QY 318 RVITIGSSDSTVAVWVNTGEMNLTLIHCEAVLHLRFNNGMMVYCSKDRSLAVMDMASP 377  
DB 283 RVITIG-SDSTVAVWVNTGEMNLTLIHCEAVLHLRFNNGMMVYCSKDRSLAVMDMASP 341  
QY 378 TDTLRRVAVGHRAAVNVDFDKYIVSASGDTIKVWNTSTCEFYRTLNHGRGIACIQ 437  
DB 342 TDTLRRVAVGHRAAVNVDFDKYIVSASGDTIKVWNTSTCEFYRTLNHGRGIACIQ 401  
QY 438 YRDLVAVSGSSDNTIRLMDIEGACLRVLEGHBEVLCIRFDNKRIVSGAYDGKIKYWDL 497  
DB 402 YRDLVAVSGSSDNTIRLMDIEGACLRVLEGHBEVLCIRFDNKRIVSGAYDGKIKYWDL 461

Related to  
Result 1

QY 498 VAALDPRAPAGTLCRLTLVHSGRVRLQDFEIVSSSHDDTLTLPNDP 550  
 Db 462 VAALDPRAPAGTLCRLTLVHSGRVRLQDFEIVSSSHDDTLTLPNDP 514

RESULT 5  
 US-09-177-165A-30  
 ; Sequence 30, Application US/09177165A  
 ; Patent No. 6426205  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tyers, Mike  
 ; APPLICANT: Williams, Andrew  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN  
 ; TITLE OF INVENTION: DEPENDENT PROTEOLYSIS  
 ; FILE REFERENCE: 11757.10USU1  
 ; CURRENT APPLICATION NUMBER: US/09/177,165A  
 ; CURRENT FILING DATE: 1998-10-22  
 ; PRIOR APPLICATION NUMBER: 60/092,443  
 ; PRIOR FILING DATE: 1998-07-10  
 ; PRIOR APPLICATION NUMBER: 60/063,254  
 ; PRIOR FILING DATE: 1997-10-24  
 ; NUMBER OF SEQ ID NOS: 50  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 30  
 ; LENGTH: 640  
 ; TYPE: PRT  
 ; ORGANISM: Saccharomyces cerevisiae  
 US-09-177-165A-30

Query Match 18.0%; Score 545; DB 4; Length 640;  
 Best Local Similarity 28.9%; Pred. No. 6e-49;  
 Matches 167; Conservative 85; Mismatches 202; Indels 124; Gaps 17;

QY 34 IPEKNSLRQTYNSCARLCLN-QETVCLASTAMKTENCVAK-TKLANGTSSMYPKGRKLS 91  
 Db 86 LPEKNTKPCYRNPDPSPPTACTYCKOLKRTQELNANIAKLPIQEOSDHIHIIKTS 145  
 QY 92 ASTEKEKELCVKYEQMSQSDQVEVEHLLISOMCHYOHCHINSYKLPMLQDFITAPAR 151  
 Db 146 NSNDKIRKL-----ILDGILSTSCFPQLSYSSLYVHMKIDPISLTP-- 188  
 QY 152 GLDHIAMNLSYDAKSIACAEVCKEYKRVYSDGMILKKLIRKRVRTISLRGLAERR- 210  
 Db 189 --GELSKILSYLDCOSLCAATKCRKWKOLADDDVWYHMCROH-----DRKC 236  
 QY 211 --GWSQYLEKFK-----PPDGNAPPNSFYRALYPRKIIDITETISNNRCGRHSQ 258  
 Db 237 PNCGWGLPILHMKRARIQONSTGSSSNADIQOTTPRWKVIYERKRVESNNRKG----- 291  
 QY 259 RIRGSESTK-----GYCLOQYDQKIVSGIRNDNTIKIMDKNTLECKRIILTGHTSVLCQ 314  
 Db 292 --HCRIOEFKHMVGVLITQFNRLFTGSDYSTIGIMDLFTGKILIRLSGHSQVKTLY 349  
 QY 315 YDERVILITSSDSTVRVWMDNGEMNLTLIHCEAVLHLRFNNGMAYVYCSKRSIAVWVM 374  
 Db 350 FDKRIITSLDKTIRVWNTITGECISTIRGSDSVLSDYOKQIVVSGSAUKYKVVWVH 409  
 QY 375 ASFTDITLRLVGVHRAAVNVYDFDKYL--VSASGDRITIKVWNTSTCEVFTLNGH-- 429  
 Db 410 ESRTCYTLR-----GHEWNVCKVLPKSPSCSCSDDTLIRMDIRTNCLAVFGRHQVQ 465  
 QY 430 KRGIACLOYRD--RLVSSSS----- 448  
 Db 466 VQKILPILTKIDENLATDNTSDGSSPDPTMTDGADESDEPSNDETVLDENIPYTHL 525  
 QY 449 -----DNTIRLMDIEGACLRVLEGEHELVRICIRPDNKRIVSGAYDKIKVMDIWAALDP 503  
 Db 526 LSGGLDNTIKIMDVYKGRKIRTOFGHVEGWMDIADNFIISSGSHDGIKAVMDLQSG--- 582  
 QY 504 KAPAGTLCRLTLVHSGRVRLQDFEIVSSSHDDTL 541  
 Db 583 -----KCHHTF---NGR---RLQRETOHTOTQSLGDKV 609

RESULT 6  
 US-08-899-578-2  
 ; Sequence 2, Application US/08899578  
 ; Patent No. 6087153  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Greenwald, Iva  
 ; APPLICANT: Hubbard, E. Jane  
 ; TITLE OF INVENTION: SEL-10 AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Cooper & Dunham LLP  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/899,578  
 ; FILING DATE: 24-JUL-1997  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: White, John P.  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 0575/53200/JPW/ABC  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 278-0400  
 ; TELEFAX: (212) 278-0525  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 587 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-899-578-2

Query Match 17.1%; Score 520; DB 3; Length 587;  
 Best Local Similarity 28.8%; Pred. No. 2.4e-46;  
 Matches 150; Conservative 78; Mismatches 222; Indels 70; Gaps 17;

QY 72 KTKLANGSSMT-----VPKQK--LSASYKEKEL-----CYKYEQMSQDVE 115  
 Db 35 ESSYSNCGSSSSYNADKLSSRPLQHKIDLSASPSRNDLNPVEHLIALFQDLSSARQMD 94  
 QY 116 FVEHLISQCHYOHCHINSYKLPMLQDFITAPALPARGLDHIAENILSYLDAKSIACAEV 175  
 Db 95 AFTRLQSSNMNTNIRQLRAITIEHPQDFLSCLPV-----ELGKILNLTGYDLKVAQV 150  
 QY 176 CKEWYRVYSDGMILKKL-IERV-----RTDLSLNGLAERRRGWGYLFRKKNPPDGNAP 227  
 Db 151 SKNKKLISEIDIKWKSGLVEEFKHHPPDPTDRVTGAMGTALAG-----VPIPDIO 203  
 QY 228 PN-SFYRALYPRKIIDITETISNNRCGRHSIQRIHCHSESTSKGYTL 273  
 Db 204 CDLVNHRFLKQKGFDFERRADKSRYLAKDKIERNNANPIMGSAV-LGHEDHVTTCM 262  
 QY 274 QYDDQKIVSGIRNDNTIKIMDKNTLECKRIILTGHTSVLCQYDE--RVITGSSDSTVRV 331  
 Db 263 QIHDDVLTGSDDDTLKAWCIDKGEVYTLVGHGTGWTSQISQCGKIVSGSDTRTKVY 322  
 QY 332 WDVNTGKRLNLTILHCEAVLHLRFNNGMAYVYCSKRSIAVWMDASPTDITLRLVGVHRA 391  
 Db 323 WSTVDSGLTLHTLQHTSTVRCMAMAGSILVYGSHTDTLIRVWDESGRHIA---TLGHHA 379  
 QY 392 AVNVVDFDDKIVASGDRITIKVWNTSTCEVFTLNGHKGKGIACLOYRD--LVVSSSSD 449  
 Db 380 AVRQVQEDGTIVVSGYDFVTYKIMNAHTGRCIRTLGHNHNVYSLLEPSRSTVCSGLD 439



Query Match 13.2%; Score 399; DB 4; Length 779;  
 Best Local Similarity 24.3%; Pred. No. 3e-33;  
 Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

QY 56 TVCLASTAKTEKNCVAKTILANGTSSMIVPKOKKLSASYEKER-----ELCVKFEQW 108  
 D 194 TPLAKTTKTINN-----NNNIADLIESKDSIIISPEYLSDEIFSAIINNMLPAAFK-- 244  
 QY 109 SESQVEFEVHEHLISQWCHYGHINSYLKPMLODFITALPARGLDHAENILSYLDKAS 168  
 D 245 -----NLFRLVAMNDRELSLDTGLIKDKNLKRLITSLPF-----EISLKFNTLOFED 294  
 QY 169 LCAAEIYCKEYR-VTSDGMLKLLI--ERMYRTDSLWGLAERBGWGYLFKNKPPDGN 225  
 D 295 IINSIGVSQNMNKIIRKSTSLMKLLISENFV-----SPKGF 331  
 QY 226 APPNSEYRALYPKIIOD-----IET--TESNMGCRHSIORHCRSETSKGYCLOYD 276  
 D 332 NSINLKLISQKPKLSOODRLRLSELENIFILKNWYKPVQRTTLRGHMTSVITCLOFE 391  
 QY 277 DQKIVSGLRDNTIKINDKNTLECKRIILGHTGSVLCLOYDE-RVYITGSSDSTVRWMDVN 335  
 D 392 DNYVITGADDKMIRYVDSINRKFLQLSGHDGVMALKYAHGILVSGSDTRIVRWMDIK 451  
 QY 336 TGEMLNTLIHCEAVLHLRFNNGMAYTCSKDRSLAVWMDASPTDITLRVLYGHRAAVNV 395  
 D 452 KG-----CCT-----HVEEGHNSIVRC 468  
 QY 396 VDDDD-----KIYVSASGDRTIKVWNT-----STGE-----FVPTLNG 428  
 D 469 LDIVYKKNIKIYVTSGRDNTLHWKLPKRESSVPDHGEHNDYPLVHTPEENPFVGYLNG 528  
 QY 429 HKRGIAQYDRDLVYSSSDNTIRLMDIEGACLRVLEGHHELVRCIRPDN--KRIYSG 486  
 D 529 HMASVFTVSGHGNIYVSSGYDNTLLVWDVAKMKCLYILSGHTRISTYIYHEKRCISA 588  
 QY 487 AYDGKIKVWL-----VAALDPRAPAGTL--CLRTLVEHSGRVRLQFDEQIYVS 535  
 D 589 SMDTITRIMDLLENIMNGECSYATNSASPCAKIIGAMTYTLOGHTALVGLRLSDKFLVSA 648  
 QY 536 SHDDTILMDPLNDPAAQAEPPRSPSRITYY 566  
 D 649 AADGSIRGMD-AND-----YSRKFSTY 668

RESULT 9  
 US-08-473-089-32  
 Sequence 32, Application US/08473089  
 Patent No. 6342368  
 GENERAL INFORMATION:  
 APPLICANT: Mochly-Rosen, Daria  
 APPLICANT: Ron, Dorit  
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
 TITLE OF INVENTION: Thereof  
 NUMBER OF SEQUENCES: 265  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Morrison & Foerster  
 STREET: 2000 Pennsylvania Avenue, NW  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20006-1812  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/473,089  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.  
 REGISTRATION NUMBER: 29,959  
 REFERENCE/DOCKET NUMBER: 2550-0025.22  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 887-1500  
 TELEFAX: (202) 887-0763  
 INFORMATION FOR SEQ ID NO: 32:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 779 amino acids  
 TYPE: amino acid  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 HYPOTHEICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 INDIVIDUAL ISOLATE: CD4 / CD20 protein, fig. 15  
 US-08-473-089-32

Query Match 13.2%; Score 399; DB 4; Length 779;  
 Best Local Similarity 24.3%; Pred. No. 3e-33;  
 Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

QY 56 TVCLASTAKTEKNCVAKTILANGTSSMIVPKOKKLSASYEKER-----ELCVKFEQW 108  
 D 194 TPLAKTTKTINN-----NNNIADLIESKDSIIISPEYLSDEIFSAIINNMLPAAFK-- 244  
 QY 109 SESQVEFEVHEHLISQWCHYGHINSYLKPMLODFITALPARGLDHAENILSYLDKAS 168  
 D 245 -----NLFRLVAMNDRELSLDTGLIKDKNLKRLITSLPF-----EISLKFNTLOFED 294  
 QY 169 LCAAEIYCKEYR-VTSDGMLKLLI--ERMYRTDSLWGLAERBGWGYLFKNKPPDGN 225  
 D 295 IINSIGVSQNMNKIIRKSTSLMKLLISENFV-----SPKGF 331  
 QY 226 APPNSEYRALYPKIIOD-----IET--TESNMGCRHSIORHCRSETSKGYCLOYD 276  
 D 332 NSINLKLISQKPKLSOODRLRLSELENIFILKNWYKPVQRTTLRGHMTSVITCLOFE 391  
 QY 277 DQKIVSGLRDNTIKINDKNTLECKRIILGHTGSVLCLOYDE-RVYITGSSDSTVRWMDVN 335  
 D 392 DNYVITGADDKMIRYVDSINRKFLQLSGHDGVMALKYAHGILVSGSDTRIVRWMDIK 451  
 QY 336 TGEMLNTLIHCEAVLHLRFNNGMAYTCSKDRSLAVWMDASPTDITLRVLYGHRAAVNV 395  
 D 452 KG-----CCT-----HVEEGHNSIVRC 468  
 QY 396 VDDDD-----KIYVSASGDRTIKVWNT-----STGE-----FVPTLNG 428  
 D 469 LDIVYKKNIKIYVTSGRDNTLHWKLPKRESSVPDHGEHNDYPLVHTPEENPFVGYLNG 528  
 QY 429 HKRGIAQYDRDLVYSSSDNTIRLMDIEGACLRVLEGHHELVRCIRPDN--KRIYSG 486  
 D 529 HMASVFTVSGHGNIYVSSGYDNTLLVWDVAKMKCLYILSGHTRISTYIYHEKRCISA 588  
 QY 487 AYDGKIKVWL-----VAALDPRAPAGTL--CLRTLVEHSGRVRLQFDEQIYVS 535  
 D 589 SMDTITRIMDLLENIMNGECSYATNSASPCAKIIGAMTYTLOGHTALVGLRLSDKFLVSA 648  
 QY 536 SHDDTILMDPLNDPAAQAEPPRSPSRITYY 566  
 D 649 AADGSIRGMD-AND-----YSRKFSTY 668

RESULT 10  
 US-08-487-072A-32  
 Sequence 32, Application US/08487072A  
 Patent No. 6423684  
 GENERAL INFORMATION:  
 APPLICANT: Mochly-Rosen, Daria  
 APPLICANT: Ron, Dorit  
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
 NUMBER OF SEQUENCES: 265

```

CORRESPONDENCE ADDRESS:
ADDRESS: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,072A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: MORASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 779 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: CDC4 / CDC20 protein, Fig. 15
US-08-487-072A-32

Query Match 13.2%; Score 399; DB 4; Length 779;
Best Local Similarity 24.3%; Pred. No. 3e-33;
Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

56 TVCLASTAMKTEVCVAKTKLANGSSMIYVKKORIKLSASYEKER-----ELCVYFEQW 108
194 TTPLAATTKTINN-----NNNIADLESKDSITISPELSDERISAINNNLPHAYFR-- 244
109 SESDQVEFEVHLISOMCHYOHGHSYLYKPMLODFITALPARGLDHIAENILSYDAKS 168
245 -----NLFLRYANNDRELSLDIGTLIKDNLKRLDITSLP-----EISLKIFNYLOFED 294
169 LCAAEVLCKEWR-VTSDGMLKKLI--ERNVRTSLMRLGLAERKGGQYLFKNKPPDGN 225
295 IINSLVGSOMNKKIIRKSTSLMKKLLISENFV-----SPKGF 331
226 APPNSFYRALYPKIIOD-----IET--IESNMRCGRHSLORIHCRSETSKEYVCLQYD 276
332 NSLNLKLSQRYKPLSGOODRLRLSFLFNITLKNMYPKFPVORTLRLGHMTSVITCLOPE 391
277 DOKIVSGLRDNTIKIMDKNTLECKRILTGHTGSVLCLOYDE-RVLIITGSSDSTVRWMDVN 335
392 DNVITIGADDMKMLRYVDSINKKFLQLSGHDGVMALKYAAGLIVSGSDTRVRYWDIK 451
336 TGEMLNTLHHCBAVLHLRFNNGMAYVCSKDRSIAVMDMASPTDITLRYLVGHRAAVNV 395
452 KG-----CCT-----HYFECHNSTVRC 468
396 VDQDD---KYIVASGDRITIKVWNT-----STCE-----FVRTLNG 428
469 LDIVEYKNIKITYTGSRDNTLHWKLPKSSVDPDGEHNDYPLVFNTPREMPYFVGLRG 528
429 HKRGIAQLQYDRLVYSGSSDNTIRLMDIEGACLRVLEGEHELVCRIEDN--KRIYSG 486
529 HMASVRYVSGHGNIVYSGSDNTLIYWDVYOMKCLYILSGHTRISTYIDHERKRCISA 588
487 AYDGKIKVMDL-----VAALDPRAPAGTL--CLRLVHSGVRFVLODFEFOIVSS 535

```

```

DB 589 SMDTIRIMDLNINWNECSYATNSASPCAKIIGAMTYTLOGHIALYGLRLSDKFLVSA 648
QY 536 SHDDTILIMDELNDPAAQAEPPRSPSRITYY 566
DB 649 AADSGIRGMD-AND-----YSRKFY 668

RESULT 11
US-09-177-165A-29
Sequence 29, Application US/09177165A
Patent No. 6426205
GENERAL INFORMATION:
APPLICANT: Tyers, Mike
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN
FILE REFERENCE: 11757.10USU1
CURRENT APPLICATION NUMBER: US/09/177,165A
PRIORITY FILING DATE: 1998-10-22
PRIORITY FILING DATE: 1998-07-10
PRIORITY APPLICATION NUMBER: 60/063,254
NUMBER OF SEQ ID NOS: 50
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 29
LENGTH: 779
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-177-165A-29

Query Match 13.2%; Score 399; DB 4; Length 779;
Best Local Similarity 24.3%; Pred. No. 3e-33;
Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

56 TVCLASTAMKTEVCVAKTKLANGSSMIYVKKORIKLSASYEKER-----ELCVYFEQW 108
194 TTPLAATTKTINN-----NNNIADLESKDSITISPELSDERISAINNNLPHAYFR-- 244
109 SESDQVEFEVHLISOMCHYOHGHSYLYKPMLODFITALPARGLDHIAENILSYDAKS 168
245 -----NLFLRYANNDRELSLDIGTLIKDNLKRLDITSLP-----EISLKIFNYLOFED 294
169 LCAAEVLCKEWR-VTSDGMLKKLI--ERNVRTSLMRLGLAERKGGQYLFKNKPPDGN 225
295 IINSLVGSOMNKKIIRKSTSLMKKLLISENFV-----SPKGF 331
226 APPNSFYRALYPKIIOD-----IET--IESNMRCGRHSLORIHCRSETSKEYVCLQYD 276
332 NSLNLKLSQRYKPLSGOODRLRLSFLFNITLKNMYPKFPVORTLRLGHMTSVITCLOPE 391
277 DOKIVSGLRDNTIKIMDKNTLECKRILTGHTGSVLCLOYDE-RVLIITGSSDSTVRWMDVN 335
392 DNVITIGADDMKMLRYVDSINKKFLQLSGHDGVMALKYAAGLIVSGSDTRVRYWDIK 451
336 TGEMLNTLHHCBAVLHLRFNNGMAYVCSKDRSIAVMDMASPTDITLRYLVGHRAAVNV 395
452 KG-----CCT-----HYFECHNSTVRC 468
396 VDQDD---KYIVASGDRITIKVWNT-----STCE-----FVRTLNG 428
469 LDIVEYKNIKITYTGSRDNTLHWKLPKSSVDPDGEHNDYPLVFNTPREMPYFVGLRG 528
429 HKRGIAQLQYDRLVYSGSSDNTIRLMDIEGACLRVLEGEHELVCRIEDN--KRIYSG 486
529 HMASVRYVSGHGNIVYSGSDNTLIYWDVYOMKCLYILSGHTRISTYIDHERKRCISA 588
487 AYDGKIKVMDL-----VAALDPRAPAGTL--CLRLVHSGVRFVLODFEFOIVSS 535
536 SHDDTILIMDELNDPAAQAEPPRSPSRITYY 566

```





Query Match 11.7%; Score 354; DB 2; Length 410;  
 Best Local Similarity 29.28; Pred. No. 6.4e-29;  
 Matches 87; Conservative 57; Mismatches 110; Indels 44; Gaps 8;

```

QY 280 IVSGIRDNTIKIMDKNTLECKRIITGHTGSLCQYDE--RVITGSSDSTVRWVDVNTG 337
    :||| ||| ||| : : ||| ||| : : ||| ||| : : ||| ||| : : ||| |||
Db 123 MVSASEDATIKVMDYETGDFERTLKSHDSDVQDISPDHSGKLASCADMTIKLMDPQGF 182
    :||| ||| ||| : : ||| ||| : : ||| ||| : : ||| ||| : : ||| |||
QY 338 EMLNTLIHCEAVLHLRF--NGMMVTCSKDRSIAVWDMASPTDITLRVYLVGHRAAVNY 395
    :||| ||| ||| : : ||| ||| : : ||| ||| : : ||| ||| : : ||| |||
Db 183 ECIRTMHGHDHNVSSVAIMPDGDIIVASRDRTIKMWEVQTYGV--KFTGCHREVRM 239
    :||| ||| ||| : : ||| ||| : : ||| ||| : : ||| ||| : : ||| |||
QY 396 V--DEDDKIIVASAGDRTIKVNTSTCEFYRTLNHKGRIACQYRDR----- 441
    :||| ||| ||| : : ||| ||| : : ||| ||| : : ||| ||| : : ||| |||
Db 240 VRPNQDSTILASCSNDQTVKVVVATRECKAEI|REH|HVVECI|SMAP|ESSYS|SEATGS 299
    :||| ||| ||| : : ||| ||| : : ||| ||| : : ||| ||| : : ||| |||
QY 442 -----LVVSSSDNTIRLMDIEGACLRVLEGHELVRCIRFDN--KRIVGAYD 489
    :||| ||| ||| : : ||| ||| : : ||| ||| : : ||| ||| : : ||| |||
Db 300 ETKKSGKPGFLLSGSRDKTIKMDVSTGCLMTLVGHDMWVRGVLFHSGKFI|LSCADD 359
    :||| ||| ||| : : ||| ||| : : ||| ||| : : ||| ||| : : ||| |||
QY 490 GKIKVMDLVVALDPRAPAGTLCRTLVHSGRVFR|Q|PDEF--QIVSSSHDITLIMD 545
    :||| ||| ||| : : ||| ||| : : ||| ||| : : ||| ||| : : ||| |||
Db 360 KTLKVM DY-----KNKRCMKTLNAHEHFTSLDFHKTAPYVVTGSDQTVKWE 408
    :||| ||| ||| : : ||| ||| : : ||| ||| : : ||| ||| : : ||| |||
    
```

Search completed: February 20, 2003, 09:56:08  
 Job time : 20 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 09:54:27 ; Search time 39 seconds  
(without alignments)  
3006.175 Million cell updates/sec

Title: US-09-601-168b-2  
Perfect score: 3034  
Sequence: 1 MDPAAVLQEKALFKMNSSE.....PAAQEPSPSPRTTYISR 569

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp-archaea:\*  
2: sp-bacteria:\*  
3: sp-fungi:\*  
4: sp-human:\*  
5: sp-invertebrate:\*  
6: sp-mammal:\*  
7: sp-mhc:\*  
8: sp-organelle:\*  
9: sp-phage:\*  
10: sp-plant:\*  
11: sp-rodent:\*  
12: sp-virus:\*  
13: sp-vertebrate:\*  
14: sp-unclassified:\*  
15: sp-virus:\*  
16: sp-bacteriophage:\*  
17: sp-archaeal:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2997	98.8	569	11	090015	O9015 mus musculu
2	2990	98.5	569	11	092159	O92159 mus musculu
3	2978	98.2	569	11	092157	O92157 mus musculu
4	2286	75.3	563	11	092380	O92380 mus musculu
5	2089.5	68.9	510	5	044382	O44382 drosophila
6	2089.5	68.9	510	5	09VDE3	O9VDE3 drosophila
7	1498	49.4	430	5	09B454	O9B454 heterodera
8	1230	40.5	252	11	0922C7	O922C7 mus musculu
9	661	21.8	1326	5	09VZF4	O9VZF4 drosophila
10	640	21.1	353	4	09N0X6	O9N0X6 homo sapien
11	640	21.1	561	4	096R12	O96R12 homo sapien
12	640	21.1	589	4	096LE0	O96LE0 homo sapien
13	640	21.1	627	4	096A16	O96A16 homo sapien
14	640	21.1	707	4	0969H0	O969H0 homo sapien
15	639	21.1	629	11	08VHP4	O8VHP4 mus musculu
16	639	21.1	629	11	08VBA4	O8VBA4 mus musculu

17	638	21.0	122	13	080UN3	O80UN3 xenopus lae
18	524	17.3	585	5	095ZT0	095ZT0 caenorhabdi
19	520	17.1	587	5	044083	044083 caenorhabdi
20	469	15.5	1376	3	08XIP2	08XIP2 podospora a
21	461.5	15.2	1356	3	08XIP4	08XIP4 podospora a
22	460.5	15.2	1356	3	08XIP5	08XIP5 podospora a
23	460.5	15.2	1356	3	08XIP3	08XIP3 podospora a
24	433.5	13.6	1227	16	0820R1	0820R1 anabaena sp
25	393	13.0	77	13	080UN8	080UN8 xenopus lae
26	387.5	12.8	304	11	0904T2	0904T2 mus musculu
27	378.5	12.5	1189	16	08YLO9	08YLO9 anabaena sp
28	375	12.4	1747	16	0820Z0	0820Z0 anabaena sp
29	365.5	12.0	942	5	096611	096611 dictyostell
30	361.5	11.9	677	16	08YVL9	08YVL9 anabaena sp
31	358.5	11.8	559	16	08YSC0	08YSC0 anabaena sp
32	357	11.8	410	13	090ZL4	090ZL4 xenopus lae
33	356	11.7	410	13	090PR5	090PR5 gallus gall
34	354	11.7	410	4	08WZ89	08WZ89 homo sapien
35	351	11.6	410	6	09GL51	09GL51 sus scrofa
36	344	11.3	411	5	096698	096698 drosophila
37	340	11.2	410	4	08WZ88	08WZ88 homo sapien
38	339.5	11.2	777	3	09USN3	09USN3 schizosacch
39	336	11.1	676	16	08YSG6	08YSG6 anabaena sp
40	335.5	11.1	353	5	08T776	08T776 branchiosto
41	334.5	11.0	283	11	093CV3	093CV3 mus musculu
42	334.5	11.0	454	4	08WV51	08WV51 homo sapien
43	334.5	11.0	454	4	09HA09	09HA09 homo sapien
44	334.5	11.0	594	11	0922B6	0922B6 mus musculu
45	334	11.0	589	16	08YV14	08YV14 anabaena sp

## ALIGNMENTS

RESULT 1  
O90015 PRELIMINARY; PRT; 569 AA.  
AC O90015;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Ubiquitin ligase FWD1 (Beta-transducin repeat containing protein)  
(F-box WD40 repeat protein 1).  
GN BTRC OR FBXW1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99199275; PubMed=10097128;  
RA Hatakeyama S., Kitagawa M., Nakayama K., Shirane M., Matsumoto M.,  
RA Hattori K., Higashi H., Nakano H., Okumura K., Onoe K., Good R.A.,  
RA Nakayama K.-i.;  
RT "Ubiquitin-dependent degradation of IkappaBalpha is mediated by a  
RT ubiquitin ligase Skp1/Cul1/F-box protein FWD1."  
RL Proc. Natl. Acad. Sci. U.S.A. 96:3859-3863(1999).  
RN (2)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99075339; PubMed=9859996;  
RA Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,  
RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;  
RT "Identification of the receptor component of the IkappaBalpha-  
RT ubiquitin ligase."  
RL Nature 396:590-594(1998).  
RN (3)  
RP SEQUENCE FROM N.A.  
RX STRAUSBERG R.;  
RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
RN (4)  
RP SEQUENCE FROM N.A.  
RX STRAUSBERG R.;  
RT MEDLINE=21601157; PubMed=11735228;

RA Maruyama S., Hatakeyama S., Nakayama K., Ishida N., Kawakami K.,  
 RA Nakayama K.-I.;  
 RT "Characterization of a Mouse Gene (Fbxw6) That Encodes a Homologue of  
 RT Ctenophaditels elegans SEL-10.";  
 RL Genomics 78:214-222 (2001).  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF081887; AAD17755.1; -;  
 DR EMBL: AF099332; AAD08701.1; -;  
 DR EMBL: BC003989; AAO03989.1; -;  
 DR EMBL: AF391190; AAL40929.1; -;  
 DR EMBL: AF391178; AAL40929.1; JOINED.  
 DR EMBL: AF391179; AAL40929.1; JOINED.  
 DR EMBL: AF391180; AAL40929.1; JOINED.  
 DR EMBL: AF391181; AAL40929.1; JOINED.  
 DR EMBL: AF391182; AAL40929.1; JOINED.  
 DR EMBL: AF391183; AAL40929.1; JOINED.  
 DR EMBL: AF391184; AAL40929.1; JOINED.  
 DR EMBL: AF391185; AAL40929.1; JOINED.  
 DR EMBL: AF391186; AAL40929.1; JOINED.  
 DR EMBL: AF391187; AAL40929.1; JOINED.  
 DR EMBL: AF391188; AAL40929.1; JOINED.  
 DR EMBL: AF391189; AAL40929.1; JOINED.  
 DR MGD: MGI:1338871; Btrc.  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW ligase; Repeat; WD repeat.  
 SQ SEQUENCE 569 AA; 65105 MM; BCTD6544815B2296 CRC64;

Query Match 98.8%; Score 2997; DB 11; Length 569;  
 Best Local Similarity 98.6%; Pred. No. 7.1e-253;  
 Matches 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MDPAEAVLQERALKFPMNSSEEDCNGEPKRIIPKNSLRQTYNSCARCLINQETVCLIA 60  
 DB 1 MDPAEAVLQERALKFPMNSSEEDCNGEPKRIIPKNSLRQTYNSCARCLINQETVCLT 60  
 QY 61 STAKTEKNCVAKTKLANGTSSMIVPKORKLSASYEKEKEICVYFQMSSEDOVEFEHL 120  
 DB 61 STAKTEKNCVAKAKLANGTSSMIVPKORKLSASYEKEKEICVYFQMSSEDOVEFEHL 120  
 QY 121 ISQCHYOHGHINSYLPMLQDRFITALPARGLDHIAENILSYLDKSLCAELVCKEY 180  
 DB 121 ISQCHYOHGHINSYLPMLQDRFITALPARGLDHIAENILSYLDKSLCAELVCKEY 180  
 QY 121 ISQCHYOHGHINSYLPMLQDRFITALPARGLDHIAENILSYLDKSLCAELVCKEY 180  
 DB 121 ISQCHYOHGHINSYLPMLQDRFITALPARGLDHIAENILSYLDKSLCAELVCKEY 180  
 QY 181 RYVSDGMLKKLIERMVRTDSIMRGLAERGWGYLFKNKPPGNAPPSFYALYPKII 240  
 DB 181 RYVSDGMLKKLIERMVRTDSIMRGLAERGWGYLFKNKPPGNAPPSFYALYPKII 240  
 QY 241 ODITETESNMRCGRHSIORIHCRSETSKGYCYCLOYDDOKIVSGLRDNTIKIMDKNTLECK 300  
 DB 241 ODITETESNMRCGRHSIORIHCRSETSKGYCYCLOYDDOKIVSGLRDNTIKIMDKNTLECK 300  
 QY 301 RILGTGTSVLCLOYDERVITIGSSDSYFVWVDVNAAGEMLNTLIHHCCEAVLHLRFNNGM 360  
 DB 301 RILGTGTSVLCLOYDERVITIGSSDSYFVWVDVNAAGEMLNTLIHHCCEAVLHLRFNNGM 360  
 QY 361 VYCSKDRSLAVWDMASPTDITLRLVYVGHRAAVNVVDEDDKIVASGDRITKVNNTSTC 420  
 DB 361 VYCSKDRSLAVWDMASPTDITLRLVYVGHRAAVNVVDEDDKIVASGDRITKVNNTSTC 420  
 QY 421 EFVRLNGHKGRIACLOYDBRLVSSGSDNTIRLMDIEGACLRVLEGEHELVRCIFRDN 480  
 DB 421 EFVRLNGHKGRIACLOYDBRLVSSGSDNTIRLMDIEGACLRVLEGEHELVRCIFRDN 480

QY 481 KRIVSAYDQKIKVMDVLAADPPAPAGTCLRTLVHSGRVRLQDFEQIVSSSHDT 540  
 DB 481 KRIVSAYDQKIKVMDVLAADPPAPAGTCLRTLVHSGRVRLQDFEQIVSSSHDT 540  
 QY 541 IILWDFLNDPAHAEPSPSRRTYIYSR 569  
 DB 541 IILWDFLNDPAHAEPSPSRRTYIYSR 569

## RESULT 2

092159 PRELIMINARY; PRT; 569 AA.

AC 092159: 01-MAY-1999 (TRENBLREL. 10, Created)  
 DT 01-MAY-1999 (TRENBLREL. 10, Last sequence update)  
 DT 01-MAY-2002 (TRENBLREL. 20, Last annotation update)  
 DE Beta-transducin repeat containing protein.  
 GN BTRC.  
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;

SEQUENCE FROM N. A.

MDLINE=99145465; PubMed=9990853;

RA Spencer E., Jiang J., Chen Z.-J.;  
 RT "Signal-Induced Ubiquitination of IkappaBalpha by the F-box protein  
 RT Slimb/Delta-TICP.";  
 RL Genes Dev. 13:284-294 (1999).

CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF112979; AAD04181.1; -;  
 DR MGD: MGI:1338871; Btrc.

DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.

SEQUENCE 569 AA; 65047 MM; BCTC7A44815BED96 CRC64;

Query Match 98.5%; Score 2990; DB 11; Length 569;  
 Best Local Similarity 98.4%; Pred. No. 2.9e-252;  
 Matches 560; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MDPAEAVLQERALKFPMNSSEEDCNGEPKRIIPKNSLRQTYNSCARCLINQETVCLIA 60  
 DB 1 MDPAEAVLQERALKFPMNSSEEDCNGEPKRIIPKNSLRQTYNSCARCLINQETVCLT 60  
 QY 61 STAKTEKNCVAKTKLANGTSSMIVPKORKLSASYEKEKEICVYFQMSSEDOVEFEHL 120  
 DB 61 STAKTEKNCVAKAKLANGTSSMIVPKORKLSASYEKEKEICVYFQMSSEDOVEFEHL 120  
 QY 121 ISQCHYOHGHINSYLPMLQDRFITALPARGLDHIAENILSYLDKSLCAELVCKEY 180  
 DB 121 ISQCHYOHGHINSYLPMLQDRFITALPARGLDHIAENILSYLDKSLCAELVCKEY 180  
 QY 121 ISQCHYOHGHINSYLPMLQDRFITALPARGLDHIAENILSYLDKSLCAELVCKEY 180  
 DB 121 ISQCHYOHGHINSYLPMLQDRFITALPARGLDHIAENILSYLDKSLCAELVCKEY 180  
 QY 181 RYVSDGMLKKLIERMVRTDSIMRGLAERGWGYLFKNKPPGNAPPSFYALYPKII 240  
 DB 181 RYVSDGMLKKLIERMVRTDSIMRGLAERGWGYLFKNKPPGNAPPSFYALYPKII 240  
 QY 241 ODITETESNMRCGRHSIORIHCRSETSKGYCYCLOYDDOKIVSGLRDNTIKIMDKNTLECK 300  
 DB 241 ODITETESNMRCGRHSIORIHCRSETSKGYCYCLOYDDOKIVSGLRDNTIKIMDKNTLECK 300  
 QY 301 RILGTGTSVLCLOYDERVITIGSSDSYFVWVDVNAAGEMLNTLIHHCCEAVLHLRFNNGM 360  
 DB 301 RILGTGTSVLCLOYDERVITIGSSDSYFVWVDVNAAGEMLNTLIHHCCEAVLHLRFNNGM 360

```

||||| 360
DB 301 RILGHTGSLCYOYDERVITIGSSDSTVRWVDVNAEMLNTLHCEAVLHLPNNM
QY 361 VTCSSKRSIAVWDMASPDITLRLVGHRAAVNVDPDDKYIVASGRTIKVWNTSTC 420
DB 361 VTCSSKRSIAVWDMASPDITLRLVGHRAAVNVDPDDKYIVASGRTIKVWNTSTC 420
QY 421 EFVRTLNHGRGIACLOYRDRLVVGSSSDNTIRLMDIEGACLRVLEHBEVLCIRFDN 480
DB 421 EFVRTLNHGRGIACLOYRDRLVVGSSSDNTIRLMDIEGACLRVLEHBEVLCIRFDN 480
QY 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVHSGRVFLQDFEFOIVSSSHDT 540
DB 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVHSGRVFLQDFEFOIVSSSHDT 540
QY 541 ILIMDFLNDPAAQAEPPSPRTTYTISR 569
DB 541 ILIMDFLNDPAAQAEPPSPRTTYTISR 569

```

## RESULT 3

09RIG7 PRELIMINARY: PRT: 569 AA.

```

AC 09RIG7, 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Beta-transducin repeat-containing protein.
GN BTRC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Winston J., Elledge S.J., Harper J.W.;
RL Submitted (DEC-1996) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AF110396; AAD41025.1; -.
DR MGD: MGI:133871; Btrc.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPROTEINRPT.
DR PRODOM: PD000018; WD40; 4.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_6.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 569 AA; 65209 MW; E6DDCAD28D551D9D CRC64;

```

Query Match 98.2%; Score 2978; DB 11; Length 569;  
 Best Local Similarity 97.9%; Pred. No. 3.2e-251;  
 Matches 557; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

```

QY 1 MDPAAVLOEAKLFKNSSEEDNNGEPPRKIIPEKNSLRQYNSCARCLINOETVCLT 60
DB 1 MDPAAVLOEAKLFKNSSEEDNNGEPPRKIIPEKNSLRQYNSCARCLINOETVCLT 60
QY 61 STAKTECVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKTFEOMSESDDVEFEVHL 120
DB 61 STAKTECVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKTFEOMSESDDVEFEVHL 120
QY 121 ISQCHYOHGHINSYKPMLOPDTTALPARGLDHIAENIISYDAKSLCAELVCKEWMY 180
DB 121 ISQCHYOHGHINSYKPMLOPDTTALPARGLDHIAENIISYDAKSLCAELVCKEWMY 180
QY 181 RVSDGMIMKRLIERMVRTDSLIMGLAERBGWGYLTKNNPPDGNAPNSFYRALYPKII 240
DB 181 RVSDGMIMKRLIERMVRTDSLIMGLAERBGWGYLTKNNPPDGNAPNSFYRALYPKII 240

```

```

DB 181 RVSDGMIMKRLIERMVRTDSLIMGLAERBGWGYLTKNNPPDGNAPNSFYRALYPKII 240
QY 241 ODIEETESNMRCGRHSIORICRSETSKGYCLOYDDKRIYSGLRDNTIKIMDNTECK 300
DB 241 ODIEETESNMRCGRHSIORICRSETSKGYCLOYDDKRIYSGLRDNTIKIMDNTECK 300
QY 301 RILGHTGSLCYOYDERVITIGSSDSTVRWVDVNAEMLNTLHCEAVLHLPNNM 360
DB 301 RILGHTGSLCYOYDERVITIGSSDSTVRWVDVNAEMLNTLHCEAVLHLPNNM 360
QY 361 VTCSSKRSIAVWDMASPDITLRLVGHRAAVNVDPDDKYIVASGRTIKVWNTSTC 420
DB 361 VTCSSKRSIAVWDMASPDITLRLVGHRAAVNVDPDDKYIVASGRTIKVWNTSTC 420
QY 421 EFVRTLNHGRGIACLOYRDRLVVGSSSDNTIRLMDIEGACLRVLEHBEVLCIRFDN 480
DB 421 EFVRTLNHGRGIACLOYRDRLVVGSSSDNTIRLMDIEGACLRVLEHBEVLCIRFDN 480
QY 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVHSGRVFLQDFEFOIVSSSHDT 540
DB 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVHSGRVFLQDFEFOIVSSSHDT 540
QY 541 ILIMDFLNDPAAQAEPPSPRTTYTISR 569
DB 541 ILIMDFLNDPAAQAEPPSPRTTYTISR 569

```

## RESULT 4

0923HO PRELIMINARY: PRT: 563 AA.

```

AC 0923HO, 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE F-box/WD40 repeat-containing protein HOS.
GN FBXW1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Bhattacharya N., Hertel J.R., Slaga T.J., Fuchs S.Y., Spiegelman V.S.;
RT "Mouse homolog of HOS (mHOS) is overexpressed in skin tumors and
RT implicated in constitutive activation of NF-kappaB."
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AY038079; AAK72095.1; -.
DR MGD: MGI:2144023; FBXW1B.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRODOM: PD000018; WD40; 4.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 563 AA; 64741 MW; 9AB562F3FF5E3496 CRC64;

```

Query Match 75.3%; Score 2286; DB 11; Length 563;  
 Best Local Similarity 75.7%; Pred. No. 8e-191;  
 Matches 440; Conservative 46; Mismatches 65; Indels 30; Gaps 7;

```

QY 1 MDPAAVLOEAKLFKNSSEEDNNGEPPRKIIPEKNSLRQYNSCARCLINOETVCLT 60
DB 1 MDPAAVLOEAKLFKNSSEEDNNGEPPRKIIPEKNSLRQYNSCARCLINOETVCLT 60
QY 56 TVCLASTA-MKTEN-----CVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKTFEOM 108
DB 56 TVCLASTA-MKTEN-----CVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKTFEOM 108
QY 45 VRLQNTSVWEDQNEDESPKKSALMQLISNCTSSVIYSRRKPSBGNYQKQKDCICITYFDQM 104
DB 45 VRLQNTSVWEDQNEDESPKKSALMQLISNCTSSVIYSRRKPSBGNYQKQKDCICITYFDQM 104

```

QY 109 SPSDVEFEVHLSOMCHYOHSHINSYLKPMLOQDPTTALPARGLDHIAENILSYLDAKS 168  
 Db 105 SPSDVEFEVHLSOMCHYOHSHINSYLKPMLOQDPTTALPARGLDHIAENILSYLDAKS 164  
 QY 169 LCAALVCKEWMRYVSDMLMKKLERVAVRTDLSLWGLAERRGGOYLFKKKPPDGNAP 228  
 Db 165 LCAALVCKEWMRYVSDMLMKKLERVAVRTDLSLWGLAERRGGOYLFKKKPPDGNAP 222  
 QY 229 NSFVYALPKIIODIETIESNMRCGRHSLOIRHCRSETSKGVYCLQYDDKQIVSGLRDNT 288  
 Db 223 NSFVYALPKIIODIETIESNMRCGRHSLOIRHCRSETSKGVYCLQYDDKQIVSGLRDNT 282  
 QY 289 IKIMDKNTLECKRITLGTGTVLQYDERVITIGSSSTYRVWDVNTGEMNLTHHCE 348  
 Db 283 IKIMDKNTLECKRITLGTGTVLQYDERVITIGSSSTYRVWDVNTGEMNLTHHCE 342  
 QY 349 AVLHLRFNNGMMVTCSSKDRSLAVWDMASPTDITLRLVYGHRAAVNVVDFDCKIVASG 408  
 Db 343 AVLHLRFNNGMMVTCSSKDRSLAVWDMASPTDITLRLVYGHRAAVNVVDFDCKIVASG 402  
 QY 409 DRTIKVMTSTCEVPTLNGHKGRIACIQYDRVLVSGSSDNTIRLMDIEGACLRVLEG 468  
 Db 403 DRTIKVMTSTCEVPTLNGHKGRIACIQYDRVLVSGSSDNTIRLMDIEGACLRVLEG 462  
 QY 469 HEEIVRCIRFENKRTIVSGAYDGKIKWMDLVVALDPRAPAGTICLRTLVHSGRVFRLQFD 528  
 Db 463 HEEIVRCIRFENKRTIVSGAYDGKIKWMDLVVALDPRAPAGTICLRTLVHSGRVFRLQFD 522  
 QY 529 EFQIVSSSHDPTLLIMDFLNDPAQAEPSPSPSTYTYISR 569  
 Db 523 EFQIVSSSHDPTLLIMDFLNDPAQAEPSPSPSTYTYISR 563  
 RESULT 5  
 044382 PRELIMINARY; PRT; 510 AA.  
 AC 044382;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE SLMB.  
 GN SLMB OR CG3412.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98121115; PubMed=9461217;  
 RA Jiang J., Struhl G.;  
 RT Regulation of the Hedgehog and Wingless signalling pathways by the F-  
 box/WD40-repeat protein Slmb.  
 RT Nature 391:493-496(1998).  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF032878; AAC38852.1;  
 DR FlyBase: FBgn0023423; slmb.  
 DR InterPro: IPR001810; F-box.  
 DR Pfam: PF000645; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR PRINTS: PRO00320; GPROTEINBPT.  
 DR PRODOM: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 510 AA: 58952 MW: DB0243D3730A5E8 CRC64:  
 Query Match 68.9%; Score 2089.5; DB 5; Length 510;

Best Local Similarity 78.3%; Pred. No. 9, 8e-174;  
 Matches 394; Conservative 48; Mismatches 54; Indels 7; Gaps 4;  
 QY 64 MKTENCAKTKLNG---SSMIVPPKORKLSA--YEKEELCYKFEWSSSDOYEVE 118  
 Db 4 MKTENCAKTKLNG---SSMIVPPKORKLSA--YEKEELCYKFEWSSSDOYEVE 118  
 QY 119 HLISOMCHYOHSHINSYLKPMLOQDPTTALPARGLDHIAENILSYLDAKS 178  
 Db 63 HLISOMCHYOHSHINSYLKPMLOQDPTTALPARGLDHIAENILSYLDAKS 172  
 QY 179 WYVYSDGMIMKRLTRMVRTDLSLWGLAERRGGOYLFKKKPPDGNAPSPYALYK 238  
 Db 123 WYVYSDGMIMKRLTRMVRTDLSLWGLAERRGGOYLFKKKPPDGNAPSPYALYK 238  
 QY 239 IYDIETIESNMRCGRHSLOIRHCRSETSKGVYCLQYDDKQIVSGLRDNTIKIMKNTLE 298  
 Db 182 IYDIETIESNMRCGRHSLOIRHCRSETSKGVYCLQYDDKQIVSGLRDNTIKIMKNTLE 298  
 QY 299 CKRILGHTGTVLQYDERVITIGSSSTYRVWDVNTGEMNLTHHCEAVLHLRFNNG 358  
 Db 242 CKRILGHTGTVLQYDERVITIGSSSTYRVWDVNTGEMNLTHHCEAVLHLRFNNG 301  
 QY 359 MMVTCSSKDRSLAVWDMASPTDITLRLVYGHRAAVNVVDFDCKIVASGDTIYVMTS 418  
 Db 302 MMVTCSSKDRSLAVWDMASPTDITLRLVYGHRAAVNVVDFDCKIVASGDTIYVMTS 361  
 QY 419 TCBEVFTLNGHKGRIACIQYDRVLVSGSSDNTIRLMDIEGACLRVLEGHEIVRCIRF 478  
 Db 362 TCBEVFTLNGHKGRIACIQYDRVLVSGSSDNTIRLMDIEGACLRVLEGHEIVRCIRF 421  
 QY 479 DNKRIYSGAYDGKIKWMDLVVALDPRAPAGTICLRTLVHSGRVFRLQFDEQIVSSSHD 538  
 Db 422 DNKRIYSGAYDGKIKWMDLVVALDPRAPAGTICLRTLVHSGRVFRLQFDEQIVSSSHD 481  
 QY 539 DPTLLIMDFLNDPAQAEPSPSPSTYTYISR 561  
 Db 482 DPTLLIMDFLNDPAQAEPSPSPSTYTYISR 504  
 RESULT 6  
 09VDE3 PRELIMINARY; PRT; 510 AA.  
 AC 09VDE3;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE SLMB protein (SLMB).  
 GN SLMB OR CG3412.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,  
 Brandon R.C., Rogers J.-H., Blazer R.G., Champe M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 Beeson K.T., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,  
 Burlingame K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 Cherry J.M., Cawley S., Dahlke C., Dawson L.B., Davies P., Davies I.,  
 de Pablos K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

[illegible]

```

QY 193 IERWARTSILWGLAERGMWQYLEKN-----PPGMAP---PNSFYRXL 236
Db 1 IEROVRNDAIWMAGLIRRGWKKYFFCSDQOATRAISQSMRPKNGEDALFEOHFYRELX 60
QY 237 PIIODIEFTESNMWCGHSLORIHCBSETEKGYCLOYDDOKIYVSGIRNDTIKIMKNT 296
Db 61 PKITPDAHQIEONMREGNRRLERIKCSGENSENKGVCLOYDDEKILISGRINDTIKIMSPS 1220
QY 297 IECRKLITGHTGSVLCLOYDERVITITGSSDSTVRWMDVNTGEMLNTLIHCEAVLHREFN 356
Db 121 IECVOTLSTGHGTSVLCLOYDNOMIMSGSSDPAIVRIMDVNTEBOLKTLIVHCEAVLHREFQ 180
QY 357 NCMWNTCSKDRSIAVMDASPTDILRLRVLVGHRAAVWVDFDDKYIYASAGDPTIKIMWN 416
Db 181 DCMILVTCSDRSIAVMDASPRDIEVRVVLVGHRAAVWVDFDIKYYIASAGDPTIKIMWN 240
QY 417 TSTCEVRLTNGHKGLACLOYRDLVYVSGSSDNTIRLMDIECGACGLVLEGGHELYRCL 476

```

Db 241 ASTCEFTVRLNGKRGKGIACLOQRDLIVSGSSDYSIRLMDIEGNCGLVLECHEIYVNCI 300  
 QY 477 RPDNRRIYSGAVDCKIKWMDLVAAIDPAPAGTLCRLTVHSGRVFLQDFEQIYSSS 536  
 Db 301 RFDNRRIYSGAVDCKIKWMDLVAAIDPAPAGTLCRLTVHSGRVFLQDFEQIYSSS 360  
 QY 537 HDTLLIMDFIN-DEPA-----CAEPPRS 561  
 Db 361 HDTLLIMDFIN-DEPA-----CAEPPRS 405  
 RESULT 8  
 Q922C7 PRELIMINARY; PRT; 252 AA.  
 ID 0922C7; PRELIMINARY; PRT; 252 AA.  
 AC 0922C7;  
 DT 01-DEC-2001 (TREMBlrel, 19, Created)  
 DT 01-DEC-2001 (TREMBlrel, 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel, 21, Last annotation update)  
 DE Similar to f-box and WD-40 domain protein 1B (Fragment).  
 GN FBXW1B.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1 SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL; BC008552; AA080552.1; -.  
 DR MGI; MGI:2144023; Fbxw1b.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 6.  
 DR PRINTS; PR00320; GPROTEINRPT.  
 DR PRODOM; PD000018; WD40; 3.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 6.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 FT NON\_TER  
 SQ SEQUENCE 252 AA; 28424 MW; F71737CBD7A9F75F CRC64;  
 Query Match 40.58; Score 1230; DB 11; Length 252;  
 Best Local Similarity 92.98; Pred. No. 3.6e-99;  
 Matches 234; Conservative 7; Mismatches 11; Indels 0; Gaps 0;  
 QY 318 RVITGSSDSTVRVWDVNTGEMTLIHHCFAVHLRFNNMGMYTCKSDSIAMVMDASP 377  
 Db 1 RVITGSSDSTVRVWDVNTGEMTLIHHCFAVHLRFNNMGMYTCKSDSIAMVMDASP 60  
 QY 378 TDTLRRLVGHRAAVNVDFDQKYIVSASGDRTIKWMNTSTGEVTLNGHKGKGIACLO 437  
 Db 61 TDTLRRLVGHRAAVNVDFDQKYIVSASGDRTIKWMNTSTGEVTLNGHKGKGIACLO 120  
 QY 438 YRDLVYSGSSDNTIRLMDIEGACLVLECHEIYVNCI RPDNRRIYSGAVDCKIKWMDL 497  
 Db 121 YRDLVYSGSSDNTIRLMDIEGACLVLECHEIYVNCI RPDNRRIYSGAVDCKIKWMDL 180  
 QY 498 VVALDPAPAGTLCRLTVHSGRVFLQDFEQIYSSSHDDTLLIMDFINDEPAQAEPP 557  
 Db 181 VVALDPAPAGTLCRLTVHSGRVFLQDFEQIYSSSHDDTLLIMDFINDEPAQAEPP 240  
 QY 558 RSPSRITTYISR 569  
 Db 241 RSPSRITTYISR 252  
 RESULT 9  
 Q9VZF4 PRELIMINARY; PRT; 1326 AA.  
 AC 09VZF4;  
 DT 01-MAY-2000 (TREMBlrel, 13, Created)  
 DT 01-MAY-2000 (TREMBlrel, 13, Last sequence update)

DT 01-JUN-2002 (TREMBlrel, 21, Last annotation update)  
 DE CG15010 protein (L2D21322p) (L2D30271p).  
 GN Aco OR CG15010.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 CX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Goez J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champs M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.C.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Bussan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Engelstiel C.C., Ferraz C., Ferrieres S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Housh K.A., Howland T.J., Wei M.-H., Ibbegan C.,  
 RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kerchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,  
 RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos A.C., Simpson M., Skupski M.P., Smith T.,  
 RA Splet E., Spirdling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svitkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassenaar D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT The genome sequence of Drosophila melanogaster.;  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nunoo J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Frise E., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
 RA Mirada A., Mungall C.J., Nunoo J., Pacle J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Celniker S.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1 SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL; AE003480; AAG22246.1; -.  
 DR EMBL; AY061300; AAL28848.1; -.  
 DR EMBL; AY075401; AAL68231.1; -.



OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RX SEQUENCE FROM N.A.  
 RA MEDLINE=2003061; PubMed=10531037;  
 RT Winston J.T., Koepf D.M., Zhu C., Ellledge S.J., Harper J.W.,  
 RL "A family of mammalian F-box proteins.";  
 RN Curr. Biol. 9:1180-1182(1999).  
 RP SEQUENCE FROM N.A.  
 RA Koepf D.M., Winston J.T., Harper J.W., Ellledge S.J.,  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AF831178; AAK60269.1; -  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR ProDom: PD000018; WD40; 2.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KM Repeat; WD repeat.  
 FT NON-TER 1  
 SQ SHOUENCE 561 AA; 63165 MW; B81C1E2206B0D8 CRC64;

Query Match 21.1%; Score 640; DB 4; Length 561;  
 Best local similarity 30.8%; Pred. No. 3.7e-47;  
 Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

14 KPNNSERDCNNGEPKRIPEKNSLQTYNSCARLCINQETVCLASTAMKTENCYAKT 73  
 DB KIDHGEVRSFSLGKPKKY-----SEYSTTGL-----VPCSA-----TPPTFGDL 66  
 QY KLANGSSMIYPRKORLSYEKEKEL--CYTFEOMSESDOVEVEHLISOMCHYOHG 131  
 DB 67 RAANGG-----QQRRTTSVQPTGLOBWLMKQMSGPKLLADLIDSCPTQVKA 121  
 QY 132 INSYLKPMLQDFITAPARGIDHTAENILSYLDAKSLCAELCKEYRVTSDGMIMK 191  
 DB 122 MQQVIEPQFORDFISLP-----KELALYVLSFLEPKDLQAQCRVWRILADNLMRE 177  
 QY 192 LTERVVRTSDSLMRGLAE-----RRCWGQYLKKNKPPGNAPNPSFYALYPKIIQDIETI 246  
 DB 178 KCKE-----EGIDEPHLIKRRK-----VIKPGFIHSPKMSAY-----IRQ--HRI 215  
 QY 247 ESNMRCGRHSLSQIRHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECKRILTGH 306  
 DB 216 DTNMRGELKSPKY-LKGHDHVTICLQPCGNRIYVSGSDNTLAKWSAYTGKCLRTLVGH 274  
 QY 307 TGSVLCLOYDERVITITGSSDSTVRYVWVNGEMLNTLIHCEAVLHLRFNNGMAYTCSKD 366  
 DB 275 TGGVSSQMRNIIISGSTDRTLKVMNAETGECIHITLYGHTSVRCMHLHEKRYVSGSRD 334  
 QY 367 RSIAMWMASTPDTLRLRVLVGHRAAVNVDPDKYIVSASGDTIYVWNTSTICEVRYTL 426  
 DB 335 ATLRYWDLDTGQCL---HYLMGHAAVAVRCVQYDGRVAVSGAYDVMKVPDPTETCLHTL 391  
 QY 427 NGHKGRIACLOYRDLVYVSSSDNTIRLMDIEGACRLVLEGHEELVRCIRFQDKRIYVSG 486  
 DB 392 QGHTNRVYSLQFDGIHVYVSGSLDTSIRYMDVETGNCIHITLHGOSLTSGMELKDNILVSG 451  
 QY 487 AYDGKIKVMDLVVALDPRAPAGTICLRTLV---EHSGRVFRLODFEFOIVSSSHDPTILL 543  
 DB 452 NADSTVKIMDIKTG-----QCLQTLQPGNKHOSAVTCLQFNKNFVITSSDDGTIVKL 502  
 QY 544 WD 545  
 DB 503 WD 504

RESULT 12  
 ID 096LEO PRELIMINARY; PRT; 589 AA.  
 AC 096LEO;  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE F-box protein SEL10.  
 GN SEL10.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Li J., Pauley A.M., Myers R.L., Shuang R., Bashler J.R., Yan R.,  
 RT Buhl A.E., Gurney M.E.,  
 RT "SEL-10 interacts with Presenilin 1, facilitates its ubiquitination,  
 and Alters A-beta production.";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL: AV008274; AAG16640.1; -  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PF00400; WD40; 7.  
 DR ProDom: PD000018; WD40; 2.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KM Repeat; WD repeat.  
 FT NON-TER 1  
 SQ SEQUENCE 589 AA; 66120 MW; 2AFB6E8A36E6E8DE CRC64;

Query Match 21.1%; Score 640; DB 4; Length 589;  
 Best local similarity 30.8%; Pred. No. 4e-47;  
 Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

14 KPNNSERDCNNGEPKRIPEKNSLQTYNSCARLCINQETVCLASTAMKTENCYAKT 73  
 DB KIDHGEVRSFSLGKPKKY-----SEYSTTGL-----VPCSA-----TPPTFGDL 66  
 QY KLANGSSMIYPRKORLSYEKEKEL--CYTFEOMSESDOVEVEHLISOMCHYOHG 131  
 DB 53 KIDHGEVRSFSLGKPKKY-----SEYSTTGL-----VPCSA-----TPPTFGDL 94  
 QY 74 KLANGSSMIYPRKORLSYEKEKEL--CYTFEOMSESDOVEVEHLISOMCHYOHG 131  
 DB 95 RAANGG-----QQRRTTSVQPTGLOBWLMKQMSGPKLLADLIDSCPTQVKA 149  
 QY 132 INSYLKPMLQDFITAPARGIDHTAENILSYLDAKSLCAELCKEYRVTSDGMIMK 191  
 DB 150 MQQVIEPQFORDFISLP-----KELALYVLSFLEPKDLQAQCRVWRILADNLMRE 205  
 QY 192 LTERVVRTSDSLMRGLAE-----RRCWGQYLKKNKPPGNAPNPSFYALYPKIIQDIETI 246  
 DB 206 KCKE-----EGIDEPHLIKRRK-----VIKPGFIHSPKMSAY-----IRQ--HRI 243  
 QY 247 ESNMRCGRHSLSQIRHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECKRILTGH 306  
 DB 244 DTNMRGELKSPKY-LKGHDHVTICLQPCGNRIYVSGSDNTLAKWSAYTGKCLRTLVGH 302  
 QY 307 TGSVLCLOYDERVITITGSSDSTVRYVWVNGEMLNTLIHCEAVLHLRFNNGMAYTCSKD 366  
 DB 303 TGGVSSQMRNIIISGSTDRTLKVMNAETGECIHITLYGHTSVRCMHLHEKRYVSGSRD 362  
 QY 367 RSIAMWMASTPDTLRLRVLVGHRAAVNVDPDKYIVSASGDTIYVWNTSTICEVRYTL 426  
 DB 363 ATLRYWDLDTGQCL---HYLMGHAAVAVRCVQYDGRVAVSGAYDVMKVPDPTETCLHTL 419  
 QY 427 NGHKGRIACLOYRDLVYVSSSDNTIRLMDIEGACRLVLEGHEELVRCIRFQDKRIYVSG 486  
 DB 420 QGHTNRVYSLQFDGIHVYVSGSLDTSIRYMDVETGNCIHITLHGOSLTSGMELKDNILVSG 479  
 QY 487 AYDGKIKVMDLVVALDPRAPAGTICLRTLV---EHSGRVFRLODFEFOIVSSSHDPTILL 543  
 DB 480 NADSTVKIMDIKTG-----QCLQTLQPGNKHOSAVTCLQFNKNFVITSSDDGTIVKL 530

```

OY 544 WD 545
DB 531 WD 532

RESULT 13
O96A16 PRELIMINARY; PRT; 627 AA.
ID 096A16:
AC 096A16:
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
DE Archipelago beta form (F-box protein FBW7).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=21449047; PubMed=11565033;
RA Moberg K.H., Bell D.W., Maher D.C., Haber D.A., Hariharan I.K.;
RT "Archipelago regulates cyclin E levels in Drosophila and is mutated in
RT human cancer cell lines.";
RL Nature 413:311-316(2001).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE=20003061; PubMed=10531037;
RA Winston J.T., Koepf D.M., Zhu C., Ellledge S.J., Harper J.W.;
RT "A family of mammalian F-box proteins.";
RL Curr. Biol. 9:1180-1182(1999).
RN 13
RP SEQUENCE FROM N.A.
RA Koepf D.M., Winston J.T., Harper J.W., Ellledge S.J.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AF411972; AAL06291.1; -.
DR EMBL: AY033553; AAK5747.1; -.
DR InterPro: IPR001810; F-box.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 2.
DR ProDom: PD000018; WD40; 2.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SO SEQUENCE 627 AA; 70324 MW; 3D4107C05381BED CRC64;

Query Match 21.1%; Score 640; DB 4; Length 627;
Best local similarity 30.8%; Pred. No. 4, 4e-47;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;
OY 14 KFMNSSEEDCNGEPKRIIPKNSLRQYNSCARLCLNOETVCLASTAMKTEVCYAKT 73
DB 91 KLDGSEVRSFSLGKKPKCV-----SEYTSITGL-----VPCSA-----TPPTFGDL 132
OY 74 KLANGTSMIVPKORKLASVKEKEKEL--CVKFEQMSSESDOYEFVHLISQCHYQGH 131
DB 133 RAAGGQC-----QQRKRITVSOPPTGLOEWKMKQSWSGPEKLALDELIDSCPTOVKH 187
OY 132 INSYLKPMLOKRDFTTALPAKGLDIAENILSYDPAKSLCAELVCKEYVTSQGLMK 191
DB 188 MMQVTEPQFQDFISLP-----KELALYVLSFLEPKDLQAQTCRWRRLAEENLIMRE 243
OY 192 LIEHNVRTDSLMKGLA-----RKGWQYLFKNKPPDGNAPNSFYALYPKIIQDIETI 246
DB 244 KCKE-----EGIDPLHTRRK-----VIKPGFHSFMSKAY-----IRQ--HRI 281
OY 247 ESNRCGSHSLDRTHCHSETSKGVYCLQYDDOKITVSGLRDITIKWKNTLECKRIITGH 306
DB 282 DTNMRGELSPKV-LKGHDDHVTTCLOFGCNRIVSGSDNTLTKWMSAVTGCKLRLVGH 340

```

```

OY 307 TGSVLCLOYDERVITITGSSDSTVRYWVNTGEMLTLLIHCEAVLHLRFNNGMAYTCSKD 366
DB 341 TGVWSSQMRDNIISGSDTFLTKVNAETGECHITLHGHTSTVACMLHREKRVSGSRD 400
OY 367 RSLAVWDMASPDITLRLRVLGHRAAVNVYDPDDKYYISASQDKRIKWNSTGCFVTL 426
DB 401 ATLRLWADIETGGCL---HLLMGHVAARVCVOYDGRVVGAYDFWVKWDEPETETCLHTL 457
OY 427 NGHKRGIACTOYRDRUVVSGSSDNTIRLMDIEGACALRVLEGHELVRCIRFDRNKRIYSG 486
DB 458 OCHTRNRYVSLQPDGTHVYVSGSLDTISRWVDYTGNCITHTLGHOSLTSGMELKDNILVSG 517
OY 487 AYDGKIKWDLVAALDPRAPACTLCLRLV---EHSGRVFLQDFEQIVSSSHDDTLLI 543
DB 518 NDSVTYKIMIDKITG-----QCLQTLQPGNKHQSVAVTCLQFKNKRVITSSDDGTATKL 568
OY 544 WD 545
DB 569 WD 570

```

```

RESULT 14
O969H0 PRELIMINARY; PRT; 707 AA.
ID 0969H0:
AC 0969H0:
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
DE F-box protein CDC4 (Archipelago alpha form).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=21449048; PubMed=11565034;
RA Strömhalder H., Spruck C.H., Kaiser P., Won K.A., Sangfelt O.,
RA Reed S.I.;
RT "Human F-box protein hcdc4 targets cyclin E for proteolysis and is
RT mutated in a breast cancer cell line.";
RL Nature 413:316-322(2001).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE=21449047; PubMed=11565033;
RA Moberg K.H., Bell D.W., Maher D.C., Haber D.A., Hariharan I.K.;
RT "Archipelago regulates cyclin E levels in Drosophila and is mutated in
RT human cancer cell lines.";
RL Nature 413:311-316(2001).
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AY049984; AAL07271.1; -.
DR EMBL: AF411971; AAL06290.1; -.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 2.
DR ProDom: PD000018; WD40; 2.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SO SEQUENCE 707 AA; 79663 MW; EA4357F76DFD8203 CRC64;

Query Match 21.1%; Score 640; DB 4; Length 707;
Best local similarity 30.8%; Pred. No. 5, 2e-47;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;
OY 14 KFMNSSEEDCNGEPKRIIPKNSLRQYNSCARLCLNOETVCLASTAMKTEVCYAKT 73
DB 171 KLDGSEVRSFSLGKKPKCV-----SEYTSITGL-----VPCSA-----TPPTFGDL 212

```

QY 74 KLANGSSMIVPORKRLSAYEKEL--CYKPFQMSQDVEVEHLLISOMCHYOGH 131  
Db 213 RAANGOG-----QQRRTTSVOPPTGLOEMLKMFQMSGPEKLALDELIDSCPTGVKH 267  
QY 132 INSYLKPMLOPDEFTALPARGLDIAENILSYLDAKSLCAELVCKEWMYRTSDGMLMK 191  
Db 268 MMQVTEPQFQDFISLIP-----KELALVYLSFLEPKDLQAQOTCRWRIIAEDNLMRE 323  
QY 192 LIEMWRTSLMRGLAE-----RRCMGQYLFKNKPPDGNAPNSFYRALYPKIIODIETI 246  
Db 324 KCKE-----EGIDEPHLIKRRK-----VTKPGFIHSPKMSAY-----IRO--HRI 361  
QY 247 ESNMRCGRHSIORICRSETSGKYCLOYDDOKIYSGLRDNTIKIMDKNTLECKRIILGH 306  
Db 362 DNNRRGELKSPKV-LKGHDHVITCLOFCGRRIYSGSDNTIKWMSAVTGKCLRTLVGH 420  
QY 307 TGSVLCLOYDERVITITSSDSTVRYWVDNTGEMTLIIHCEAVLHLRFNNGMMVYTSKSD 366  
Db 421 TGGVWSSQMRDNIISGSTDTLTKVWNAETGECIHTLYGHTSTVRCMHLHEKRVYSGSRD 480  
QY 367 RSLAVWMASTPDTILRLRYLVGHRRAAVNVDPDDKXIYASGDRITIKWNTSTCEPVRTL 426  
Db 481 ATLWMDIETGQCL--HYLMGHVAAVRCVQYDGRVYSGAYDFMKVWDEPTECTLHTL 537  
QY 427 NGHKGRIACLOYRDLRYVSSSDNTIRLMDIEGACILRVLEGEHELVRCLRFNPKRIYSG 486  
Db 538 QGHTRNRYVSLQFDGIHVYSGSLDTSIRWVDYETGNCIHTLGHOSLISGMLKONILVSG 597  
QY 487 AYDGKIKVMDLVALDPRAPAGTLCRLTY---EHSGRVFLRQDEFOIVSSSHDDTILI 543  
Db 598 NADSTVKIMDIKTG-----QCLQTLQPGPKHOSAVTCLQFKNKVFITSSDDGTIVKL 648  
QY 544 WD 545  
Db 649 WD 650

RESULT 15  
Q8VHP4  
ID Q8VHP4 PRELIMINARY; PRT; 629 AA.  
AC Q8VHP4;  
DT 01-MAR-2002 (Tremblrel. 20, Created)  
DT 01-MAR-2002 (Tremblrel. 20, last sequence update)  
DE 01-JUN-2002 (Tremblrel. 21, last annotation update)  
DE F-box protein.  
GN FBWD6.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Ilyin G.P.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF427101; AAL50052.1; -  
DR InterPro: IPR001810; F-box.  
DR InterPro: IPR001880; WD40.  
DR Pfam: PF00646; F-box; 1.  
DR PRINTS: PRO0320; GPROTEINRPT.  
DR ProDom: PD000018; WD40; 2.  
DR SMART: SM00256; FBOX; 1.  
DR SMART: SM00320; WD40; 7.  
DR PROSITE: PS00181; FBOX; 1.  
DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_5.  
DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
SQ SEQUENCE 629 AA; 70562 MW; BE916405A3490A3E CRC64;

Query Match 21.1%; Score 639; DB 11; Length 629;  
Best Local Similarity 30.8%; Pred. No. 5,4e-47;  
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

QY 14 KEMNSEREDCNGNEPPRKIIPEKNSLRQTYNSCARCLINDETVCLASTAMKTENCVAKT 73  
Db 93 KLDHGSEVASFISIGKKPKV-----SDYRSTTGL-----VPCSA-----TPTTFDDL 134  
QY 74 KLANGSSMIVPORKRLSAYEKEL--CYKPFQMSQDVEVEHLLISOMCHYOGH 131  
Db 135 RAANGOG-----QQRRTTSVOPPTGLOEMLKMFQMSGPEKLALDELIDSCPTGVKH 189  
QY 132 INSYLKPMLOPDEFTALPARGLDIAENILSYLDAKSLCAELVCKEWMYRTSDGMLMK 191  
Db 190 MMQVTEPQFQDFISLIP-----KELALVYLSFLEPKDLQAQOTCRWRIIAEDNLMRE 245  
QY 192 LIEMWRTSLMRGLAE-----RRCMGQYLFKNKPPDGNAPNSFYRALYPKIIODIETI 246  
Db 246 KCKE-----EGIDEPHLIKRRK-----VTKPGFIHSPKMSAY-----IRO--HRI 283  
QY 247 ESNMRCGRHSIORICRSETSGKYCLOYDDOKIYSGLRDNTIKIMDKNTLECKRIILGH 306  
Db 284 DNNRRGELKSPKV-LKGHDHVITCLOFCGRRIYSGSDNTIKWMSAVTGKCLRTLVGH 342  
QY 307 TGSVLCLOYDERVITITSSDSTVRYWVDNTGEMTLIIHCEAVLHLRFNNGMMVYTSKSD 366  
Db 343 TGGVWSSQMRDNIISGSTDTLTKVWNAETGECIHTLYGHTSTVRCMHLHEKRVYSGSRD 402  
QY 367 RSLAVWMASTPDTILRLRYLVGHRRAAVNVDPDDKXIYASGDRITIKWNTSTCEPVRTL 426  
Db 403 ATLWMDIETGQCL--HYLMGHVAAVRCVQYDGRVYSGAYDFMKVWDEPTECTLHTL 459  
QY 427 NGHKGRIACLOYRDLRYVSSSDNTIRLMDIEGACILRVLEGEHELVRCLRFNPKRIYSG 486  
Db 460 QGHTRNRYVSLQFDGIHVYSGSLDTSIRWVDYETGNCIHTLGHOSLISGMLKONILVSG 519  
QY 487 AYDGKIKVMDLVALDPRAPAGTLCRLTY---EHSGRVFLRQDEFOIVSSSHDDTILI 543  
Db 520 NADSTVKIMDIKTG-----QCLQTLQPGPKHOSAVTCLQFKNKVFITSSDDGTIVKL 570  
QY 544 WD 545  
Db 571 WD 572

Search completed: February 20, 2003, 09:55:19  
Job time : 44 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 09:58:07 : Search time 14 Seconds  
(without alignments)  
1038.379 Million cell updates/sec

Title: US-09-601-168b-2  
Perfect score: 3034  
Sequence: 1 MDPARAVIQEAKLKFMSNSE.....PAAQAEPPSPSRITYISR 569

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 140259 seqs, 2554876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications -AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCIT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCITUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3034	100.0	569	12	US-10-042-417-2
2	1163	38.3	219	9	US-10-023-530-2
3	989.5	32.6	265	10	US-09-764-848-30
4	644.5	21.2	626	9	US-09-213-888-21
5	644.5	21.2	626	9	US-09-328-877A-21
6	640	21.1	540	9	US-09-213-888-7
7	640	21.1	540	9	US-09-213-888-10
8	640	21.1	540	9	US-09-328-877A-7
9	640	21.1	540	9	US-09-328-877A-10
10	640	21.1	545	9	US-09-213-888-6
11	640	21.1	545	9	US-09-328-877A-6
12	640	21.1	553	9	US-09-213-888-5
13	640	21.1	553	9	US-09-328-877A-5
14	640	21.1	559	9	US-09-213-888-9
15	640	21.1	559	9	US-09-328-877A-9
16	640	21.1	589	9	US-09-213-888-8
17	640	21.1	589	9	US-09-328-877A-8
18	640	21.1	592	9	US-09-213-888-4
19	640	21.1	592	9	US-09-328-877A-4

20	640	21.1	627	9	US-09-213-888-3	Sequence 3, Appl1
21	640	21.1	627	9	US-09-328-877A-3	Sequence 2, Appl1
22	640	21.1	666	9	US-09-213-888-27	Sequence 27, Appl1
23	640	21.1	666	9	US-09-328-877A-27	Sequence 27, Appl1
24	640	21.1	669	9	US-09-213-888-25	Sequence 25, Appl1
25	640	21.1	669	9	US-09-328-877A-25	Sequence 25, Appl1
26	575	19.0	678	10	US-09-801-368-314	Sequence 314, App
27	545	18.0	678	10	US-10-060-019-30	Sequence 30, Appl1
28	455.5	15.0	1356	9	US-10-077-111-10	Sequence 10, Appl1
29	399	13.2	779	9	US-10-060-019-29	Sequence 29, Appl1
30	392	12.9	732	10	US-09-994-485-8	Sequence 8, Appl1
31	384	12.7	732	10	US-09-832-292-12	Sequence 12, Appl1
32	337	11.1	1146	9	US-09-994-485-6	Sequence 6, Appl1
33	337	11.1	1146	9	US-10-132-744A-2	Sequence 4, Appl1
34	334.5	11.0	422	12	US-10-042-417-4	Sequence 6, Appl1
35	316.5	10.4	485	9	US-10-132-744A-6	Sequence 6, Appl1
36	307.5	10.1	742	9	US-10-077-111-11	Sequence 11, Appl1
37	299.5	9.9	1194	10	US-09-876-667-2	Sequence 2, Appl1
38	299.5	9.9	1205	10	US-09-876-667-16	Sequence 16, Appl1
39	296.5	9.8	484	9	US-10-132-744A-2	Sequence 2, Appl1
40	293.5	9.7	261	9	US-10-132-744A-4	Sequence 4, Appl1
41	283	9.3	713	10	US-09-801-368-408	Sequence 408, App
42	264.5	8.7	423	10	US-09-729-674-160	Sequence 160, App
43	260.5	8.6	375	9	US-10-119-932-1	Sequence 1, Appl1
44	254.5	8.4	316	10	US-09-828-310-12	Sequence 12, Appl1
45	246.5	8.1	343	9	US-10-119-932-5	Sequence 5, Appl1

## ALIGNMENTS

RESULT 1	US-10-042-417-2	Application US/10042417
Sequence 2, Appl1	US-10-042-417-2	US-10-042-417-2
Patent No.	US20020123082A1	
GENERAL INFORMATION:		
APPLICANT:	Parano, M.	
TITLE OF INVENTION:	METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS	
FILE REFERENCE:	5914-090-999	
CURRENT APPLICATION NUMBER:	US/10/042,417	
CURRENT FILING DATE:	2002-01-07	
PRIOR APPLICATION NUMBER:	60/260,179	
PRIOR FILING DATE:	2001-01-5	
NUMBER OF SEQ ID NOS:	89	
SOFTWARE:	PatentIn Ver. 2.0	
SEQ ID NO 2		
LENGTH:	569	
TYPE:	PRT	
ORGANISM:	Homo sapiens	
US-10-042-417-2		
Query Match	100.0%;	Score 3034; DB 12; Length 569;
Best Local Similarity	100.0%;	Pred. No. 6.5e-248;
Matches	569; Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	1 MDPARAVIQEAKLKFMSNSE.....PAAQAEPPSPSRITYISR 60	Sequence 2, Appl1
DB	1 MDPARAVIQEAKLKFMSNSE.....PAAQAEPPSPSRITYISR 60	Sequence 2, Appl1
QY	61 STAMKTEKVAKTLANSTSSMIVPKQKLSASYEKEKELCVYKFEQWSSDQVEVEHL 120	Sequence 2, Appl1
DB	61 STAMKTEKVAKTLANSTSSMIVPKQKLSASYEKEKELCVYKFEQWSSDQVEVEHL 120	Sequence 2, Appl1
QY	121 ISOMCHYGHINSLYKPLMDLQDFITLALPANGLDHIAENILSYLDASLCAAEIVCKEWY 160	Sequence 160, App
DB	121 ISOMCHYGHINSLYKPLMDLQDFITLALPANGLDHIAENILSYLDASLCAAEIVCKEWY 160	Sequence 160, App
QY	181 RYTSOGMLMKKLTIERMVRDLSMRGLAERGGVLFNKKPPDGNAPNSFYALYKII 240	Sequence 240, Appl1
DB	181 RYTSOGMLMKKLTIERMVRDLSMRGLAERGGVLFNKKPPDGNAPNSFYALYKII 240	Sequence 240, Appl1
QY	241 ODIEETIESNMGRHSIORHCRSETSKGVYCLQYDQK IVSGLRDTIKITWKDNTLECK 300	Sequence 300, Appl1
DB	241 ODIEETIESNMGRHSIORHCRSETSKGVYCLQYDQK IVSGLRDTIKITWKDNTLECK 300	Sequence 300, Appl1

```

Db 241 QDIETESNMWRCGRHSLORHCRSETSKGVYCLQYDQKIVSGRDQTKITMKNLTCECK 300
QY 301 RLTGHSVYCLQYDEVVITGSSDSTVWVNDVNTGEMLTNLHHCVAVLHLPNNNGM 360
Db 301 RLTGHSVYCLQYDEVVITGSSDSTVWVNDVNTGEMLTNLHHCVAVLHLPNNNGM 360
QY 361 VTGSKRSIAVWMASTPDLIRLVGLHRAAVVDFDKYIVSAGDPTVWNTSTG 420
Db 361 VTGSKRSIAVWMASTPDLIRLVGLHRAAVVDFDKYIVSAGDPTVWNTSTG 420
QY 421 EFVTLNGHKGKIGACLOYRDRLVYSSSDNTIRLMDIEGACLRVLEGEHELVCIRFDN 480
Db 421 EFVTLNGHKGKIGACLOYRDRLVYSSSDNTIRLMDIEGACLRVLEGEHELVCIRFDN 480
QY 481 KRISGAVDGKIKYWDIVLADPRAPAGTLCRLTVEHSGRVRLQDFEFOIVSSSHDT 540
Db 481 KRISGAVDGKIKYWDIVLADPRAPAGTLCRLTVEHSGRVRLQDFEFOIVSSSHDT 540
QY 541 ILTMDPLNDPAQAEPSPSRRTTYISR 569
Db 541 ILTMDPLNDPAQAEPSPSRRTTYISR 569

```

## RESULT 2

```

US-10-023-530-2
; Sequence 2, Application US/10023530
; Publication No. US20030007956A1
; GENERAL INFORMATION:
; APPLICANT: LEGRAIN, Pierre
; APPLICANT: BENAROUS, Richard
; APPLICANT: BLOT, Guillaume
; APPLICANT: LASSOT, Irina
; TITLE OF INVENTION: PROTEINS THAT INTERACT WITH BETA TRCP
; FILE REFERENCE: B4717A
; CURRENT APPLICATION NUMBER: US/10/023,530
; CURRENT FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 60/256,276
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 219
; TYPE: PRP
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: BetaTrcp
; LOCATION: (1)..(219)
; OTHER INFORMATION: F-box protein
US-10-023-530-2

```

```

Query Match
Best Local Similarity 38.3%; Score 1163; DB 9; Length 219;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPAAVLQERKALFKMNSEREDCNNGEPKRIIPKNSLRQYNSCARLINOETVCLA 60
Db 1 MDPAAVLQERKALFKMNSEREDCNNGEPKRIIPKNSLRQYNSCARLINOETVCLA 60
QY 61 STAMTEKNCVAKTILANGTSSMIVPKOKLSASYEKELCVKFFEQWSESDVEVEH 120
Db 61 STAMTEKNCVAKTILANGTSSMIVPKOKLSASYEKELCVKFFEQWSESDVEVEH 120
QY 121 ISOMCHYOHGHSINYLKPMLODRFTLPAKGIDHIAENILSYLDAKSLCAAEVCKEY 180
Db 121 ISOMCHYOHGHSINYLKPMLODRFTLPAKGIDHIAENILSYLDAKSLCAAEVCKEY 180
QY 181 RYTSOGMLMKKLIRWVYRTDSLWGLAERGWGYLFKN 219
Db 181 RYTSOGMLMKKLIRWVYRTDSLWGLAERGWGYLFKN 219

```

## RESULT 3

```

US-09-764-848-30
; Sequence 30, Application US/09764848
; Patent No. US20020077270A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT208
; CURRENT APPLICATION NUMBER: US/09/764,848
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 265
; TYPE: PRP
; ORGANISM: Homo sapiens
US-09-764-848-30

```

## Query Match

```

Best Local Similarity 32.6%; Score 989.5; DB 10; Length 265;
Matches 186; Conservative 24; Mismatches 21; Indels 39; Gaps 3;

```

```

QY 15 FMNSEREDCNNGEPKRIIPKNSLRQYNSCARLINOETVCLA STAMTEKNCVAKTK 74
Db 20 FQNTSVMEQDNEDESPK-----KNTLMO----- 42
QY 75 LANGTSSMIVPKOKLSASYEKELCVKFFEQWSESDVEVEHLSOMCHYOHGHS 134
Db 43 ISNGTSVIVSRKRPSEGYOKKEDLCIKYFQWSESDVEVEHLSOMCHYOHGHS 102
QY 135 YLKPMLQDRFTLPAKGIDHIAENILSYLDAKSLCAAEVCKEYRYSOGMLMKLIE 194
Db 103 YLKPMLQDRFTLPAKGIDHIAENILSYLDAKSLCAAEVCKEYRYSOGMLMKLIE 162
QY 195 RMYRTDSLWGLAERGWGYLFKNKPPGNAAPNSFRLAYPKTIIDTIESNMWRCGR 254
Db 163 RMYRTDSLWGLAERGWGYLFKNKPPGNAAPNSFRLAYPKTIIDTIESNMWRCGR 220
QY 255 HSLQRIHCRSEFSKGVYCLQYDQKIVSG 284
Db 221 HSLQRIHCRSEFSKGVYCLQYDQKIVSG 250

```

## RESULT 4

```

US-09-213-888-21
; Sequence 21, Application US/09213888A
; Patent No. US2002016463A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; TITLE OF INVENTION: Encode Them
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/213,888A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 626
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 6 myc tagged
; OTHER INFORMATION: homo sapien
US-09-213-888-21

```

## Query Match

```

Best Local Similarity 21.2%; Score 644.5; DB 9; Length 626;
Matches 169; Conservative 103; Mismatches 222; Indels 67; Gaps 15;

```

```

QY 8 LQERKALFKMNSEREDCNNGEPKRIIPK--NSLRQ--TYNSCARLINOETVCLA STA 63

```

```

Db 53 MEOKLISEEDLMESEGLTMEOKLISEEDLNSMKRKLDSHSEVFSLSGKPPCKVSEY 112
QY 64 MKENCV-----AKTKLANGTSSMIVPORKLSASYEKEL--CVKFEQWSESD 112
Db 113 TSTGTGLVCSAIPPTTGGDLRAANGOG-----QORRRTTSVOPPTGLQEWLKMFOSSNGPE 167
QY 113 OVEFEVHLISOMCHQOHINSYLKPMLOPDTALPARGLDHAENILSYLDASISCAA 172
Db 168 KLALDELIDSCPTQVHMMQVIEPQFORDISLP---KELALVLSLEPKDLLOA 223
QY 173 ELVCEKMYRVTSDGMLMKLIERMVRTDSLWRGLAE-----RRCMGQYLFKNKPPDGNAP 227
Db 224 AQTCRYWRILAEADNLIMRECKE-----EGIDELPHIKRRK-----VIRKPGFIHSP 269
QY 228 PMSFYALPKLIODIETIESNMRCGRHSIORHCRSETSKGYVCLQYDOKIVSGLRDN 287
Db 270 WKSAY-----IRQ--HRIDTNRKGLKSPKV--LKGDHDDVITTCLOFCGNRIYVSGSDN 320
QY 288 TIKINDKNTLECKRILTGHTGSVLCLOYDERVITITGSSDSTVRVWDVNTGEMLNTLIHC 347
Db 321 TLKWSAVTGVKCLRTLVGHTGCVWSSQMRDNIISGSTDRTLKVMNAETGECIHTLYGHT 380
QY 348 EAVLHLRFNNGMMVYCSKDRSIAYVDMASPTDITLRLVYGHRAAVNVYDFDCKIYVAS 407
Db 381 STVRCHMLHEKRYVSGSRDALTFRVMDIETGCL--HYLMGHVAARVQYDGRVYVSGA 437
QY 408 GDRITKVMNTSCFVRVLNKGKRGIAQLQYRDLVYSGSSDNTIRLMDIEGACRLYLE 467
Db 438 YDFMVKVMDPETETCLHTLQGHTRVYSLQDFGIHVYSGSLDTSIRVMDVETGNCIHTLT 497
QY 468 GHEELVRCIRFDRNKRIYSGAYDKIKVMDVLAALDPRAPAGTLCRLTY--EHSGRVFR 524
Db 498 GHOSLTSGMELKDNILVSGNADSTVKIMDIKTG-----QCLQTLQGPKNKHQSAVTC 548
QY 525 LQDFEQIVSSSHDDTILMD 545
Db 549 LQFNKNFVITSSDDGTVKLMD 569

```

## RESULT 5

```

US-09-328-877A-21
; Sequence 21, Application US/09328877A
; Patent No. US2002017187A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/328,877A
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 6 myc tagged
; OTHER INFORMATION: homo sapien
US-09-328-877A-21

```

## Query Match

21.1%; Score 644.5; DB 9; Length 626;  
 Best Local Similarity 30.1%; Pred. No. 2.1e-46;  
 Matches 169; Conservative 103; Mismatches 222; Indels 67; Gaps 15;

```

QY 8 LQEALAFNMSSEREDCNNGEPPRKIIPEK--NSLRQ--TYNSCARLICINQETVCIASSTA 63
Db 53 MEOKLISEEDLMESEGLTMEOKLISEEDLNSMKRKLDSHSEVFSLSGKPPCKVSEY 112

```

```

QY 64 MKENCV-----AKTKLANGTSSMIVPORKLSASYEKEL--CVKFEQWSESD 112
Db 113 TSTGTGLVCSAIPPTTGGDLRAANGOG-----QORRRTTSVOPPTGLQEWLKMFOSSNGPE 167
QY 113 OVEFEVHLISOMCHQOHINSYLKPMLOPDTALPARGLDHAENILSYLDASISCAA 172
Db 168 KLALDELIDSCPTQVHMMQVIEPQFORDISLP---KELALVLSLEPKDLLOA 223
QY 173 ELVCEKMYRVTSDGMLMKLIERMVRTDSLWRGLAE-----RRCMGQYLFKNKPPDGNAP 227
Db 224 AQTCRYWRILAEADNLIMRECKE-----EGIDELPHIKRRK-----VIRKPGFIHSP 269
QY 228 PMSFYALPKLIODIETIESNMRCGRHSIORHCRSETSKGYVCLQYDOKIVSGLRDN 287
Db 270 WKSAY-----IRQ--HRIDTNRKGLKSPKV--LKGDHDDVITTCLOFCGNRIYVSGSDN 320
QY 288 TIKINDKNTLECKRILTGHTGSVLCLOYDERVITITGSSDSTVRVWDVNTGEMLNTLIHC 347
Db 321 TLKWSAVTGVKCLRTLVGHTGCVWSSQMRDNIISGSTDRTLKVMNAETGECIHTLYGHT 380
QY 348 EAVLHLRFNNGMMVYCSKDRSIAYVDMASPTDITLRLVYGHRAAVNVYDFDCKIYVAS 407
Db 381 STVRCHMLHEKRYVSGSRDALTFRVMDIETGCL--HYLMGHVAARVQYDGRVYVSGA 437
QY 408 GDRITKVMNTSCFVRVLNKGKRGIAQLQYRDLVYSGSSDNTIRLMDIEGACRLYLE 467
Db 438 YDFMVKVMDPETETCLHTLQGHTRVYSLQDFGIHVYSGSLDTSIRVMDVETGNCIHTLT 497
QY 468 GHEELVRCIRFDRNKRIYSGAYDKIKVMDVLAALDPRAPAGTLCRLTY--EHSGRVFR 524
Db 498 GHOSLTSGMELKDNILVSGNADSTVKIMDIKTG-----QCLQTLQGPKNKHQSAVTC 548
QY 525 LQDFEQIVSSSHDDTILMD 545
Db 549 LQFNKNFVITSSDDGTVKLMD 569

```

## RESULT 6

```

US-09-213-888-7
; Sequence 7, Application US/09213888A
; Patent No. US20020164683A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Encode Them
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/213,888A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-213-888-7

```

## Query Match

21.1%; Score 640; DB 9; Length 540;  
 Best Local Similarity 30.8%; Pred. No. 4.1e-46;  
 Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

```

QY 14 KFMNSEREDCNNGEPPRKIIPEKNSLRQTYNSCARLICINQETVCIASSTAMKTEVCVAKT 73
Db 4 KIDHGEVAFSFLGKKPKCV-----SEYTSITGL-----VPCSA-----TPTTFDGL 45
QY 74 KLANGTSSMIVPORKLSASYEKEL--CVKFEQWSESDOVEFEVHLISOMCHQOH 131
Db 46 RAANGOG-----QORRRTTSVOPPTGLQEWLKMFOSSNGPEKILALDELIDSCPTQV 100
QY 132 INSYLKPMLQRPDTLTPARGLDHAENILSYLDASISCAAALVYCKEYRVTSDGMLMKR 191

```

```

Db 101 MMQVLEPQFQDFSLP-----KELALYVLSLEPDKDLAQAOTCRYWRIIAEDNLMRE 156
QY 192 LIERNVRLDSLMRGIAE-----RGWGQYLFKNRPPDGNAPNSFYALYPIIIDIETI 246
Db 157 KCKE-----EGIDELHAKRRK-----VIRPGFIHSWKSAY-----IRQ--HRI 194
QY 247 ESNMRCGRHSLORIHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIMKNTLECKRIITGH 306
Db 195 DTNMRGELKSPKV-LKCHDDHVTICLQFCGNRIYVSGSDNTILKWSAAYGKCLRTLYGH 253
QY 307 TGSVLCLOYDERVLIITGSSDSTVAVDNTGEMLNTLIHCEAVLHLRFNNGMAYTCSKD 366
Db 254 TGVWSSQMRDNIITIGSDTRTLKWNNAETGECIHTLYGHTSTVRCMHLHEKRYVSGSRD 313
QY 367 RSIAYWDMASPTDITLRLVYGHRAAVNVDPDDKIVASAGDRTIKWNTSTCEVFTL 426
Db 314 ATLWMDIEFGQCL---HYLMGHVAAYRCVOYDGRVYSGAYDPMVKWMDPETETCLHTL 370
QY 427 NGHRGIACTQYRDLRYVSGSSDNTIRLMDIEGACLRLEHEELVRCIRFDNKRIVSG 486
Db 371 OGHTRNVYSIQFDGIHVYSGSLDTSIRWMDVETGNCIHTLTHGSLTSGMELKDNILVSG 430
QY 487 AYDGKIKVMDVLAALDPRAPAGTLCRTLY---EHSGRVRLQDFEPIYVSSSDHTILI 543
Db 431 NADSTVKIMDIKTG-----QCLQTLQGNPKHQSAYVTCLOFNKNFVITSSDDGTIVKL 481
QY 544 WD 545
Db 482 WD 483

```

```

RESULT 7
US-09-213-888-10
; Sequence 10, Application US/09213888A
; Patent No. US2002016483A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/213,888A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-213-888-10

```

```

Query Match 21.1%; Score 640; DB 9; Length 540;
Best Local Similarity 30.8%; Pred. No. 4.1e-46;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

```

```

QY 14 KFMNSREDCCNNGEPPRKIIPEKNSLRQTYNSCARLCLNGETVCLASTAMKTENCYAKT 73
Db 4 KLDHSEVRSFSLGKKPKCV-----SEYTSITGL-----VPCSA-----TPTFGDL 45
QY 74 KLANGTSSMIVPORKLISYEKEKEL--CVKFEQWSESQDVEFEVHLISOMCHYOHGH 131
Db 46 RAANGOG-----QQRRTITSVQPTGLQEWLKMFGQSMGPKKLALDELIDSCPTGYKH 100
QY 132 INSYLKPMLQDFITALPARGLDHAENILSYDAKSICAAELVCKEYRYVTSDGMLMKR 191
Db 101 MMQVLEPQFQDFSLP-----KELALYVLSLEPDKDLAQAOTCRYWRIIAEDNLMRE 156
QY 192 LIERNVRLDSLMRGIAE-----RGWGQYLFKNRPPDGNAPNSFYALYPIIIDIETI 246
Db 157 KCKE-----EGIDELHAKRRK-----VIRPGFIHSWKSAY-----IRQ--HRI 194

```

```

QY 247 ESNMRCGRHSLORIHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIMKNTLECKRIITGH 306
Db 195 DTNMRGELKSPKV-LKCHDDHVTICLQFCGNRIYVSGSDNTILKWSAAYGKCLRTLYGH 253
QY 307 TGSVLCLOYDERVLIITGSSDSTVAVDNTGEMLNTLIHCEAVLHLRFNNGMAYTCSKD 366
Db 254 TGVWSSQMRDNIITIGSDTRTLKWNNAETGECIHTLYGHTSTVRCMHLHEKRYVSGSRD 313
QY 367 RSIAYWDMASPTDITLRLVYGHRAAVNVDPDDKIVASAGDRTIKWNTSTCEVFTL 426
Db 314 ATLWMDIEFGQCL---HYLMGHVAAYRCVOYDGRVYSGAYDPMVKWMDPETETCLHTL 370
QY 427 NGHRGIACTQYRDLRYVSGSSDNTIRLMDIEGACLRLEHEELVRCIRFDNKRIVSG 486
Db 371 OGHTRNVYSIQFDGIHVYSGSLDTSIRWMDVETGNCIHTLTHGSLTSGMELKDNILVSG 430
QY 487 AYDGKIKVMDVLAALDPRAPAGTLCRTLY---EHSGRVRLQDFEPIYVSSSDHTILI 543
Db 431 NADSTVKIMDIKTG-----QCLQTLQGNPKHQSAYVTCLOFNKNFVITSSDDGTIVKL 481
QY 544 WD 545
Db 482 WD 483

```

```

RESULT 8
US-09-328-877A-7
; Sequence 7, Application US/09328877A
; Patent No. US20020177187A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/328,877A
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 7
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-328-877A-7

```

```

Query Match 21.1%; Score 640; DB 9; Length 540;
Best Local Similarity 30.8%; Pred. No. 4.1e-46;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

```

```

QY 14 KFMNSREDCCNNGEPPRKIIPEKNSLRQTYNSCARLCLNGETVCLASTAMKTENCYAKT 73
Db 4 KLDHSEVRSFSLGKKPKCV-----SEYTSITGL-----VPCSA-----TPTFGDL 45
QY 74 KLANGTSSMIVPORKLISYEKEKEL--CVKFEQWSESQDVEFEVHLISOMCHYOHGH 131
Db 46 RAANGOG-----QQRRTITSVQPTGLQEWLKMFGQSMGPKKLALDELIDSCPTGYKH 100
QY 132 INSYLKPMLQDFITALPARGLDHAENILSYDAKSICAAELVCKEYRYVTSDGMLMKR 191
Db 101 MMQVLEPQFQDFSLP-----KELALYVLSLEPDKDLAQAOTCRYWRIIAEDNLMRE 156
QY 192 LIERNVRLDSLMRGIAE-----RGWGQYLFKNRPPDGNAPNSFYALYPIIIDIETI 246
Db 157 KCKE-----EGIDELHAKRRK-----VIRPGFIHSWKSAY-----IRQ--HRI 194
QY 247 ESNMRCGRHSLORIHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIMKNTLECKRIITGH 306
Db 195 DTNMRGELKSPKV-LKCHDDHVTICLQFCGNRIYVSGSDNTILKWSAAYGKCLRTLYGH 253
QY 307 TGSVLCLOYDERVLIITGSSDSTVAVDNTGEMLNTLIHCEAVLHLRFNNGMAYTCSKD 366

```

Db 254 TGVWSSQMRDNIISGSDRTILKLVMAETGECIHTLYGHTYRVCMLHKKRRVSSGRD 313  
 QY 367 RSLAVWDMASPDITLRLVGLVHRAAVNVVDFDDIYISASGDRITKWNSTGCEFRTL 426  
 Db 314 ATLWMDITGQCL---HVLKMGHVAARVQYDGRVVSAGAYDFWVKWMDPETETCLHTL 370  
 QY 427 NGHKGACIQLYRDLRVVSSGSDNTIRLMDIEGACLVLEGEHVLVRCIFPDNKRIVSG 486  
 Db 371 OGHTRNVYSLOPFGIHVSGSLDTSIRWMDVETGNCIHTLGHOSLTSGMELKDNILVSG 430  
 QY 487 AYDGKIKWMDVLAALDPRAPAGTLCRLTV---EHSGRVFLQDFEFOYSSSHDDTLII 543  
 Db 431 NADSTVKIMWDITG-----OCLQTLQGNKQSAVTCLOFKNKNVITSSDQGVKL 481  
 QY 544 WD 545  
 Db 482 WD 483

RESULT 9  
 US-09-328-877A-10  
 : Sequence 10, Application US/09328877A  
 : Patent No. US20020177187A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Gurney, Mark E.  
 : APPLICANT: Li, Jinhe  
 : APPLICANT: Pauley, Adele M.  
 : APPLICANT: Pharmacia & Upjohn Company  
 : TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
 : FILE REFERENCE: 6142  
 : CURRENT APPLICATION NUMBER: US/09/328,877A  
 : CURRENT FILING DATE: 1999-06-09  
 : NUMBER OF SEQ ID NOS: 27  
 : SOFTWARE: Patentln Ver. 2.0  
 : SEQ ID NO 10  
 : LENGTH: 540  
 : TYPE: PRP  
 : ORGANISM: Homo sapiens  
 : US-09-328-877A-10

Query Match 21.1%; Score 640; DB 9; Length 540;  
 Best Local Similarity 30.8%; Pred. No. 4.1e-46;  
 Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

QY 14 KFMNSEREDCNNGEPPKIIPEKNSLRQTYNSCARCLNQETVCLASTAMKTEHCYAKT 73  
 Db 4 KLDHGEVRSFSLKPKCKV-----SEYTSITGL-----VPCSA-----TPYTFGDL 45  
 QY 74 KLANGTSMIYVPOKRLSASYEKEKEL--CYKYPEQMSSEDQVEFEVHLISOMCHVOGH 131  
 Db 46 RAANGOG-----QQRRTTSVQPTGLQEWLKNFOSMSGPEKLLALDELIDSCPTQVKH 100  
 QY 132 INSYLKPMLQDRFTALPARGLDIAENIISYLDKSLCAAEVLCKEMRYVTSGLMLWK 191  
 Db 101 MMQVLEPQORDFISLP---KELALYVLSFLEPKDLQAAQTCRYWRILAEIDNLLMRE 156  
 QY 192 LIERVARTDSLMRGLAE-----RGMGOYLFPKNKPPDGNAPNSFYALYPKIIIDDIETI 246  
 Db 157 KCKE-----EGIDELPLHKRRK-----VIRGFHSHWKSAY-----IRQ--HRI 194  
 QY 247 ESNMRCGRHSIORHICRSETSGKYVCLQYDDOKIYVGLNDNTIKIMDKNTLECKRIITGH 306  
 Db 195 DTNMRGELKSPKV-LKGDHDVITCLOFCGNRIYSGSDDTLTKWMSAVTGKCLRTLVGH 253  
 QY 307 TGSVLCLOYDERVITGSSDSTVRWVDNTEMLNTLIHCEAVLHLRPNNGMMVYTCSD 366  
 Db 254 TGVWSSQMRDNIISGSDRTILKLVMAETGECIHTLYGHTYRVCMLHKKRRVSSGRD 313  
 QY 367 RSLAVWDMASPDITLRLVGLVHRAAVNVVDFDDIYISASGDRITKWNSTGCEFRTL 426  
 Db 314 ATLWMDITGQCL---HVLKMGHVAARVQYDGRVVSAGAYDFWVKWMDPETETCLHTL 370

QY 427 NGHKGACIQLYRDLRVVSSGSDNTIRLMDIEGACLVLEGEHVLVRCIFPDNKRIVSG 486  
 Db 371 OGHTRNVYSLOPFGIHVSGSLDTSIRWMDVETGNCIHTLGHOSLTSGMELKDNILVSG 430  
 QY 487 AYDGKIKWMDVLAALDPRAPAGTLCRLTV---EHSGRVFLQDFEFOYSSSHDDTLII 543  
 Db 431 NADSTVKIMWDITG-----OCLQTLQGNKQSAVTCLOFKNKNVITSSDQGVKL 481  
 QY 544 WD 545  
 Db 482 WD 483

RESULT 10  
 US-09-213-888-6  
 : Sequence 6, Application US/09213888A  
 : Patent No. US20020164683A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Gurney, Mark E.  
 : APPLICANT: Li, Jinhe  
 : APPLICANT: Pauley, Adele M.  
 : APPLICANT: Pharmacia & Upjohn Company  
 : TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
 : FILE REFERENCE: 6142  
 : CURRENT APPLICATION NUMBER: US/09/213,888A  
 : CURRENT FILING DATE: 1998-12-17  
 : NUMBER OF SEQ ID NOS: 27  
 : SOFTWARE: Patentln Ver. 2.0  
 : SEQ ID NO 6  
 : LENGTH: 545  
 : TYPE: PRP  
 : ORGANISM: Homo sapiens  
 : US-09-213-888-6

Query Match 21.1%; Score 640; DB 9; Length 545;  
 Best Local Similarity 30.8%; Pred. No. 4.2e-46;  
 Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

QY 14 KFMNSEREDCNNGEPPKIIPEKNSLRQTYNSCARCLNQETVCLASTAMKTEHCYAKT 73  
 Db 9 KLDHGEVRSFSLKPKCKV-----SEYTSITGL-----VPCSA-----TPYTFGDL 50  
 QY 74 KLANGTSMIYVPOKRLSASYEKEKEL--CYKYPEQMSSEDQVEFEVHLISOMCHVOGH 131  
 Db 51 RAANGOG-----QQRRTTSVQPTGLQEWLKNFOSMSGPEKLLALDELIDSCPTQVKH 105  
 QY 132 INSYLKPMLQDRFTALPARGLDIAENIISYLDKSLCAAEVLCKEMRYVTSGLMLWK 191  
 Db 106 MMQVLEPQORDFISLP---KELALYVLSFLEPKDLQAAQTCRYWRILAEIDNLLMRE 161  
 QY 192 LIERVARTDSLMRGLAE-----RGMGOYLFPKNKPPDGNAPNSFYALYPKIIIDDIETI 246  
 Db 162 KCKE-----EGIDELPLHKRRK-----VIRGFHSHWKSAY-----IRQ--HRI 199  
 QY 247 ESNMRCGRHSIORHICRSETSGKYVCLQYDDOKIYVGLNDNTIKIMDKNTLECKRIITGH 306  
 Db 200 DTNMRGELKSPKV-LKGDHDVITCLOFCGNRIYSGSDDTLTKWMSAVTGKCLRTLVGH 258  
 QY 307 TGSVLCLOYDERVITGSSDSTVRWVDNTEMLNTLIHCEAVLHLRPNNGMMVYTCSD 366  
 Db 259 TGVWSSQMRDNIISGSDRTILKLVMAETGECIHTLYGHTYRVCMLHKKRRVSSGRD 318  
 QY 367 RSLAVWDMASPDITLRLVGLVHRAAVNVVDFDDIYISASGDRITKWNSTGCEFRTL 426  
 Db 319 ATLWMDITGQCL---HVLKMGHVAARVQYDGRVVSAGAYDFWVKWMDPETETCLHTL 375  
 QY 427 NGHKGACIQLYRDLRVVSSGSDNTIRLMDIEGACLVLEGEHVLVRCIFPDNKRIVSG 486  
 Db 376 OGHTRNVYSLOPFGIHVSGSLDTSIRWMDVETGNCIHTLGHOSLTSGMELKDNILVSG 435  
 QY 487 AYDGKIKWMDVLAALDPRAPAGTLCRLTV---EHSGRVFLQDFEFOYSSSHDDTLII 543

Db 436 NADSTVKIMDIKTG-----QCLQTLQGPKNKHSAVTCLOFGNKNFVITSSDGTIVKL 486  
 QY 544 WD 545  
 Db 487 WD 488

RESULT 11  
 US-09-328-877A-6  
 ; Sequence 6, Application US/09328877A  
 ; Patent No. US20020177187A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gurney, Mark E.  
 ; APPLICANT: Li, Jinhe  
 ; APPLICANT: Pauley, Adele M.  
 ; APPLICANT: Pharmacia & Upjohn Company  
 ; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
 ; FILE REFERENCE: 6142  
 ; CURRENT APPLICATION NUMBER: US/09/328, 877A  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 6  
 ; LENGTH: 545  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-328-877A-6

Query Match 21.1%; Score 640; DB 9; Length 545;  
 Best Local Similarity 30.8%; Pred. No. 4.2e-46;  
 Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

QY 14 KMNSSEREDCNNGEPKRIPEKNSLRQYNSCARLQNEVCLASTAMKTEVCYAKT 73  
 Db 9 KLDHSEVRSSISLGGKPKCV-----SEYTSITGL-----VPCSA-----TPITFGDL 50  
 QY 74 KLANSTSMIYPKOKKLSASYEKEKEL--CYKFEQMSSEDOVEFEHLISQCHQOHH 131  
 Db 51 RAANGOG-----QQRRTTSVQPTGLQEWLKMFOGWSGPEKLLAIDELIDSCHEPTOVKH 105  
 QY 132 INSYLKPMLQDFETALPARGLDHIAENILSYLDAKSLCAELVCKEYRVYSDGMLMK 191  
 Db 106 MMQVTEPQFQDFISLTP-----KELALYVLSFLEPKDLQAOTCRWRILAEIDLME 161  
 QY 192 LIERNVRTDSLWRGLAE-----RRWGQOYLFKKKPPDGNAPPNPSFYRALYPKTIODIETI 246  
 Db 162 KCKE-----EGIDELPHIKRRK-----VIKPGFIHSPKMSAV-----IRQ--HRI 199  
 QY 247 ESNMRCGRHSIORHCHSETSKGYVCLQYDQKIVSGLRDNTIKIDKNTLECKRILTGH 306  
 Db 200 DTNMRGELKSPKV-LKGHDDHVTICLOFGCNRIYVSGSDNTLKVMSAVTGKCLRTLVGH 258  
 QY 307 TGSVLCLOYDERVITIGSSDSYRVMDVNTGEMLNTLIHHCNAVLIHREFNGMAYTCSKD 366  
 Db 259 TGGVWSSQMRDNIISGSTDRITLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVYSGSD 318  
 QY 367 RSLAVWMASTPTDITLRLVYGHRAAVNVPDDKIVYASGDRITKVNSTICEFVRL 426  
 Db 319 ATLRYWDIETGQCL---HYLMGHVAAVRCVQYDGRVYSGAIDFVWKWDPETETCLHTL 375  
 QY 427 NGHKGRIACLOYDRILVYSGSSDNTIRLMDIEGACRLVLEHHELVYRCIRDNKRIYSG 486  
 Db 376 OGHNTNRYYSIQDFGIHVYSGSLDTSIRVMDVETGNCIHTLYGHQSLTSGMELKMDIIVSG 435  
 QY 487 AYDGKIKVMDLVAALDPRAPAGTICLRLTV---EHSGRFRLQDFEFOIVSSSHDITLI 543  
 Db 436 NADSTVKIMDIKTG-----QCLQTLQGPKNKHSAVTCLOFGNKNFVITSSDGTIVKL 486  
 QY 544 WD 545  
 Db 487 WD 488

RESULT 12  
 US-09-213-888-5  
 ; Sequence 5, Application US/09213888A  
 ; Patent No. US20020164683A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gurney, Mark E.  
 ; APPLICANT: Li, Jinhe  
 ; APPLICANT: Pauley, Adele M.  
 ; APPLICANT: Pharmacia & Upjohn Company  
 ; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that  
 ; FILE REFERENCE: 6142  
 ; CURRENT APPLICATION NUMBER: US/09/213, 888A  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 5  
 ; LENGTH: 553  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-213-888-5

Query Match 21.1%; Score 640; DB 9; Length 553;  
 Best Local Similarity 30.8%; Pred. No. 4.3e-46;  
 Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

QY 14 KMNSSEREDCNNGEPKRIPEKNSLRQYNSCARLQNEVCLASTAMKTEVCYAKT 73  
 Db 17 KLDHSEVRSSISLGGKPKCV-----SEYTSITGL-----VPCSA-----TPITFGDL 58  
 QY 74 KLANSTSMIYPKOKKLSASYEKEKEL--CYKFEQMSSEDOVEFEHLISQCHQOHH 131  
 Db 59 RAANGOG-----QQRRTTSVQPTGLQEWLKMFOGWSGPEKLLAIDELIDSCHEPTOVKH 113  
 QY 132 INSYLKPMLQDFETALPARGLDHIAENILSYLDAKSLCAELVCKEYRVYSDGMLMK 191  
 Db 114 MMQVTEPQFQDFISLTP-----KELALYVLSFLEPKDLQAOTCRWRILAEIDLME 169  
 QY 192 LIERNVRTDSLWRGLAE-----RRWGQOYLFKKKPPDGNAPPNPSFYRALYPKTIODIETI 246  
 Db 170 KCKE-----EGIDELPHIKRRK-----VIKPGFIHSPKMSAV-----IRQ--HRI 207  
 QY 247 ESNMRCGRHSIORHCHSETSKGYVCLQYDQKIVSGLRDNTIKIDKNTLECKRILTGH 306  
 Db 208 DTNMRGELKSPKV-LKGHDDHVTICLOFGCNRIYVSGSDNTLKVMSAVTGKCLRTLVGH 266  
 QY 307 TGSVLCLOYDERVITIGSSDSYRVMDVNTGEMLNTLIHHCNAVLIHREFNGMAYTCSKD 366  
 Db 267 TGGVWSSQMRDNIISGSTDRITLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVYSGSD 326  
 QY 367 RSLAVWMASTPTDITLRLVYGHRAAVNVPDDKIVYASGDRITKVNSTICEFVRL 426  
 Db 327 ATLRYWDIETGQCL---HYLMGHVAAVRCVQYDGRVYSGAIDFVWKWDPETETCLHTL 383  
 QY 427 NGHKGRIACLOYDRILVYSGSSDNTIRLMDIEGACRLVLEHHELVYRCIRDNKRIYSG 486  
 Db 384 OGHNTNRYYSIQDFGIHVYSGSLDTSIRVMDVETGNCIHTLYGHQSLTSGMELKMDIIVSG 443  
 QY 487 AYDGKIKVMDLVAALDPRAPAGTICLRLTV---EHSGRFRLQDFEFOIVSSSHDITLI 543  
 Db 444 NADSTVKIMDIKTG-----QCLQTLQGPKNKHSAVTCLOFGNKNFVITSSDGTIVKL 494  
 QY 544 WD 545  
 Db 495 WD 496

RESULT 13  
 US-09-328-877A-5  
 ; Sequence 5, Application US/09328877A  
 ; Patent No. US20020177187A1  
 ; GENERAL INFORMATION:



LENGTH: 559  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 us-09-328-877A-9

Query Match 21.1% Score 640; DB 9; Length 559;  
 Best Local Similarity 30.8%; Pred. No. 4.3e-46;  
 Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

```

QY 14 KFNSSEREDCNNGEPPRKIIPEKNSLRQTYNSCARLCLNOETWCLASTAMKTENCYAKT 73
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 23 KLDHGEVRSFSLGKKPKCKV-----SEXTSTGL-----VPCSA-----TPTFGDL 64
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 74 KLANGTSMIVPKQKRLKASAEKEKL--CYKTPQMSSESDQVEFVHLLSQMCHYQGH 131
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 65 RAANGQG-----QQRRTTSVQPPGIG/DEWLKMFQSWSGPEKLLADELIDSEPTQYKH 119
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 132 INSLKPMLOPDTTALPARGLDHIAENILSYLDAKSLCAAEIVCKEMRYVTSQMLMK 191
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 120 MMQYIEPQFQDFISLP-----KELALYVLSFLEPKDLLQAQTCRYWRILAEEDNLMRE 175
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 192 LIERNVRTDSLWGLAE-----RRGQOYLEKKNRPDCMAPPNSFYRALYPKIIQDIET 246
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 176 KCKE-----EGIDPELHKRRK-----VIKPGFIHSPWKSAY-----IRQ--HRI 213
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 247 ESNMRCGRHSLQRIHCRSETSKGYCLOYDDOKIVSGLRDNITIKIMDKNTLECKRIITGH 306
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 214 DTNNRGEELASPKV-LKGHDHVTTCLOFGCNRIVSGSDNTLKWMSAVTGKCLRTLVGH 272
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 307 TGSVLCLOYDERVLIITGSSDSTVREWDVNTGEMLNTLIHCEAVLHLRFNNGMAYTCSKD 366
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 273 TGVWSSQMRDNIIITSGSTDRTELKWNNAETGECIHTLYGHTSTYRCMHLHEKRYVSGSRD 332
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 367 RSTAVMDASTPTDITLRKVLVGHRAAVNVDEDDKIYASASGRTIKVWNTSTCEEVRTL 426
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 333 ATLEWVDIETGQCL--HYLMGHVAARVQYDGRRYVSGAYDFMVKVMDPETETCLHTL 389
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 427 NGHRRGIACLOQRDLVYSSGSDNTIRLMDIEGACLRVLEGHELVRCIRFDKRIYVG 486
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 390 OGHTRRYVSLQFDGIHYVSSGSDTIRVWDYETGNCIHTLTGHOSLTSGMELKDNIIYVG 449
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 487 AYDGKIVWDLVAALDPRAPAGTLCRLTV--EHSGRVRLQDFEFQIVSSSHDPTILI 543
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 450 NADSTVKIMDIKTG-----OCLQTLQGPKNHQSAVTCLOPFKNKFVITSSDDGTVKL 500
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 544 WD 545
   |  |
Db 501 WD 502

```

Search completed: February 20, 2003, 10:00:10  
 Job time : 17 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 09:56:12 : Search time 57 seconds

(without alignments)  
897.640 Million cell updates/sec

Title: US-09-601-168B-2

Perfect score: 3034

Sequence: 1 MDPRAVLOEKALKFNMSSEREDCNNGEPRIKILPEKNSLRQYNSCARLCAQLQETVCLLA

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 468380 seqs, 89921735 residues

Total number of hits satisfying chosen parameters: 468380

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length DB	ID	Description
1	3034	100.0	569	US-09-601-168B-2	Sequence 2, Appl 1
2	3006	99.1	605	US-09-724-676-52731	Sequence 52731, A
3	3006	99.1	605	US-09-724-676-52731	Sequence 52731, A
4	2701	89.0	537	US-09-724-676-52729	Sequence 52729, A
5	2701	89.0	537	US-09-724-676-52730	Sequence 52730, A
6	2701	89.0	537	US-09-724-676-52729	Sequence 52729, A
7	2701	89.0	537	US-09-724-676-52730	Sequence 52730, A
8	2673	88.1	573	US-09-724-676-52732	Sequence 52732, A
9	2673	88.1	573	US-09-724-676-52732	Sequence 52732, A
10	2673	88.1	573	US-09-724-676-52733	Sequence 52733, A
11	2673	88.1	573	US-09-724-676-52733	Sequence 52733, A
12	2400	79.1	529	US-09-724-676-61015	Sequence 61015, A
13	2400	79.1	529	US-09-724-676-61016	Sequence 61016, A
14	2400	79.1	529	US-09-724-676-61015	Sequence 61015, A
15	2400	79.1	529	US-09-724-676-61016	Sequence 61016, A
16	2384.5	78.6	542	US-09-724-676-61014	Sequence 61014, A
17	2384.5	78.6	542	US-09-724-676-61014	Sequence 61014, A
18	2358	77.7	512	US-09-724-676-61017	Sequence 61017, A
19	2358	77.7	512	US-09-724-676-61018	Sequence 61018, A
20	2358	77.7	512	US-09-724-676-61017	Sequence 61017, A
21	2358	77.7	512	US-09-724-676-61018	Sequence 61018, A
22	545	18.0	640	US-10-060-019-30	Sequence 30, Appl 1
23	399	13.2	779	US-10-060-019-29	Sequence 29, Appl 1
24	354	11.7	410	US-09-724-676-61377	Sequence 61377, A
25	354	11.7	410	US-09-724-676-61378	Sequence 61378, A
26	354	11.7	410	US-09-724-676-61379	Sequence 61379, A

27	354	11.7	410	US-09-724-676-61377	Sequence 61377, A
28	354	11.7	410	US-09-724-676-61378	Sequence 61378, A
29	354	11.7	410	US-09-724-676-61379	Sequence 61379, A
30	334.5	11.0	486	US-09-724-676-60984	Sequence 60984, A
31	334.5	11.0	486	US-09-724-676-60984	Sequence 60984, A
32	334.5	11.0	600	US-10-197-666A-120	Sequence 120, App
33	334.5	11.0	631	US-10-197-666A-116	Sequence 116, App
34	334.5	11.0	667	US-10-197-666A-118	Sequence 118, App
35	334.5	11.0	670	US-10-197-666A-114	Sequence 114, App
36	332.5	11.0	357	US-10-264-237-2768	Sequence 2768, App
37	330.5	10.9	631	US-10-197-666A-144	Sequence 144, App
38	330.5	10.9	641	US-10-197-666A-112	Sequence 112, App
39	330.5	10.9	670	US-10-197-666A-150	Sequence 150, App
40	330.5	10.9	670	US-10-218-140-4360	Sequence 4360, App
41	324.5	10.7	334	PCT-US02-27671A-2	Sequence 2, Appl 1
42	324.5	10.7	334	PCT-US02-27671A-4	Sequence 4, Appl 1
43	314	10.3	524	PCT-US02-40925-3221	Sequence 3221, App
44	314	10.3	524	US-10-320-797-3221	Sequence 3221, App
45	303.5	10.0	1213	US-09-724-676-50571	Sequence 50571, A

## ALIGNMENTS

RESULT 1  
US-09-601-168B-2  
Sequence 2, Application US/09601168B  
GENERAL INFORMATION:  
APPLICANT: BENAROUS, Richard  
APPLICANT: MARGOTIN, Florence  
APPLICANT: DURAND, Hervé  
APPLICANT: AREZVANA SEIDEDOS, Fernando  
APPLICANT: KROLL, Mathias  
APPLICANT: CONDORET, Jean-Paul  
TITLE OF INVENTION: Human beta-TrCP protein  
FILE REFERENCE: 935.38812X00  
CURRENT APPLICATION NUMBER: US/09/601,168B  
CURRENT FILING DATE: 2000-07-28  
PRIOR APPLICATION NUMBER: PCT/FR99/00196  
PRIOR FILING DATE: 1999-01-29  
PRIOR APPLICATION NUMBER: FR98 01100  
PRIOR FILING DATE: 1998-01-30  
PRIOR APPLICATION NUMBER: FR98 15545  
PRIOR FILING DATE: 1998-12-09  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn Ver. 2.1 and manually  
SEQ ID NO 2  
LENGTH: 569  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: Description of the artificial sequence : CDNA  
OTHER INFORMATION: coding for human beta-TrCP protein  
US-09-601-168B-2

Query Match: 100.0% Score 3034; DB 5; Length 569;  
Best Local Similarity 100.0%; Pred No. 5e-273;  
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MDPRAVLOEKALKFNMSSEREDCNNGEPRIKILPEKNSLRQYNSCARLCAQLQETVCLLA 60  
1 MDPRAVLOEKALKFNMSSEREDCNNGEPRIKILPEKNSLRQYNSCARLCAQLQETVCLLA 60  
61 STAKTENCYAKTKRLANGTSSMIVPKOKISASYEKEKELCYFFQWSSDVEFEVHL 120  
61 STAKTENCYAKTKRLANGTSSMIVPKOKISASYEKEKELCYFFQWSSDVEFEVHL 120  
121 ISQCHQHGHSINSYLPMLORDFITLPAFGDHIENILSYLDKSLCAELVCKEY 180  
121 ISQCHQHGHSINSYLPMLORDFITLPAFGDHIENILSYLDKSLCAELVCKEY 180  
121 ISQCHQHGHSINSYLPMLORDFITLPAFGDHIENILSYLDKSLCAELVCKEY 180  
181 RVTSQMIWKKLIERVVRTSLWGLAERGGQGYLFKNPDPGNAAPNSFYALPKII 240  
181 RVTSQMIWKKLIERVVRTSLWGLAERGGQGYLFKNPDPGNAAPNSFYALPKII 240

Db 181 RVISDGMIMKKLIERVNTDSLMRGIAERRGQYLFFKNKPPDGNAPNSFYALYKII 240  
 QY 241 QDIETIESNNRCGRHSIORIHCSSETSKEYCYCLOYDDOKIVSGLRDNNTIKIMDKNTLECK 300  
 Db 241 QDIETIESNNRCGRHSIORIHCSSETSKEYCYCLOYDDOKIVSGLRDNNTIKIMDKNTLECK 300  
 QY 301 RILGHTGVSVCLOYDERVITITSSSDSYAVWVNTGEMLNTLIHCEAVLHLRFNGM 360  
 Db 301 RILGHTGVSVCLOYDERVITITSSSDSYAVWVNTGEMLNTLIHCEAVLHLRFNGM 360  
 QY 361 VTCSKDSIAVMDMASPTDITLRRVIVGHRAAVNVDPDDKTYVASGRTIKVWNTSIC 420  
 Db 361 VTCSKDSIAVMDMASPTDITLRRVIVGHRAAVNVDPDDKTYVASGRTIKVWNTSIC 420  
 QY 421 EFVRTLNGHRGIAICLOYRDLVSSGSDNTYRLMDIEGACILVLEGEHELVRICIRPDN 480  
 Db 421 EFVRTLNGHRGIAICLOYRDLVSSGSDNTYRLMDIEGACILVLEGEHELVRICIRPDN 480  
 QY 481 KRIYSGAYDGKIKVMDLVALDPRAPAGTLCITLVESHGVRFLQDFEQIVSSSHDT 540  
 Db 481 KRIYSGAYDGKIKVMDLVALDPRAPAGTLCITLVESHGVRFLQDFEQIVSSSHDT 540  
 QY 541 ILIMDFLNDPAQAEPSPRSRTYITISR 569  
 Db 541 ILIMDFLNDPAQAEPSPRSRTYITISR 569

## RESULT 2

US-09-724-676-52731  
 ; Sequence 52731, Application US/09724676  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Comugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Comugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 52731  
 ; LENGTH: 605  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-724-676-52731

Query Match 99.1%; Score 3006; DB 5; Length 605;  
 Best Local Similarity 94.0%; Pred. No. 2.2e-270;

Matches 569; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 MDPAEAVILOEKALFKM-----NSSEREDC 24  
 Db 1 MDPAEAVILOEKALFKMCMSPRSLMWGSSSLADSMPSLRCLNPGTALTFQNSSEREDC 60  
 QY 25 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTENCYAKTKLANGTSSMIV 84  
 Db 61 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTENCYAKTKLANGTSSMIV 120  
 QY 85 PRQKRLSASYEKEKELCVKFEQWSESDQVEFVHLLISQCHYOHGHINSYLKPMLODF 144  
 Db 121 PRQKRLSASYEKEKELCVKFEQWSESDQVEFVHLLISQCHYOHGHINSYLKPMLODF 180  
 QY 145 ITALPARGLDHAENILSYLDAKSLCAELVCKEMRYVTSDDGMIMKKLIERVNTDSLNR 204  
 Db 181 ITALPARGLDHAENILSYLDAKSLCAELVCKEMRYVTSDDGMIMKKLIERVNTDSLNR 240  
 QY 205 GLAERRGGOYLFFKNKPPDGNAPNSFYRALYKIIODIETIESNNRCGRHSIORIHCS 264  
 Db 241 GLAERRGGOYLFFKNKPPDGNAPNSFYRALYKIIODIETIESNNRCGRHSIORIHCS 300  
 QY 265 ETSKGVYCLQYDDOKIYVSGLDNNTIKIMDKNTLECKRILGHTGVSVCLOYDERVITITSS 324  
 Db 301 ETSKGVYCLQYDDOKIYVSGLDNNTIKIMDKNTLECKRILGHTGVSVCLOYDERVITITSS 360  
 QY 325 SDSYIRVWDVNTGEMLNTLIHCEAVLHLRFNNGMVTGSKDRIAVWDMASPTDITLRR 384

Db 361 SDSYIRVWDVNTGEMLNTLIHCEAVLHLRFNNGMVTGSKDRIAVWDMASPTDITLRR 420  
 QY 385 VLVGHRAAVNVVDPDDKTYVASGDRITIKVWNTSTCEPRTLNGHRGIAICLOYDRILV 444  
 Db 421 VLVGHRAAVNVVDPDDKTYVASGDRITIKVWNTSTCEPRTLNGHRGIAICLOYDRILV 480  
 QY 445 SGSSDNTYRLMDIEGACILVLEGEHELVRICIRPDNKRIVSGAYDGKIKVMDLVALDPR 504  
 Db 481 SGSSDNTYRLMDIEGACILVLEGEHELVRICIRPDNKRIVSGAYDGKIKVMDLVALDPR 540  
 QY 505 APAGTLCITLVESHGVRFLQDFEQIVSSSHDTILIMDFLNDPAQAEPSPRSRTY 564  
 Db 541 APAGTLCITLVESHGVRFLQDFEQIVSSSHDTILIMDFLNDPAQAEPSPRSRTY 600  
 QY 565 TYISR 569  
 Db 601 TYISR 605

## RESULT 3

US-09-724-676A-52731  
 ; Sequence 52731, Application US/09724676A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Comugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Comugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676A  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 52731  
 ; LENGTH: 605  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-724-676A-52731

Query Match 99.1%; Score 3006; DB 5; Length 605;  
 Best Local Similarity 94.0%; Pred. No. 2.2e-270;

Matches 569; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 MDPAEAVILOEKALFKM-----NSSEREDC 24  
 Db 1 MDPAEAVILOEKALFKMCMSPRSLMWGSSSLADSMPSLRCLNPGTALTFQNSSEREDC 60  
 QY 25 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTENCYAKTKLANGTSSMIV 84  
 Db 61 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTENCYAKTKLANGTSSMIV 120  
 QY 85 PRQKRLSASYEKEKELCVKFEQWSESDQVEFVHLLISQCHYOHGHINSYLKPMLODF 144  
 Db 121 PRQKRLSASYEKEKELCVKFEQWSESDQVEFVHLLISQCHYOHGHINSYLKPMLODF 180  
 QY 145 ITALPARGLDHAENILSYLDAKSLCAELVCKEMRYVTSDDGMIMKKLIERVNTDSLNR 204  
 Db 181 ITALPARGLDHAENILSYLDAKSLCAELVCKEMRYVTSDDGMIMKKLIERVNTDSLNR 240  
 QY 205 GLAERRGGOYLFFKNKPPDGNAPNSFYRALYKIIODIETIESNNRCGRHSIORIHCS 264  
 Db 241 GLAERRGGOYLFFKNKPPDGNAPNSFYRALYKIIODIETIESNNRCGRHSIORIHCS 300  
 QY 265 ETSKGVYCLQYDDOKIYVSGLDNNTIKIMDKNTLECKRILGHTGVSVCLOYDERVITITSS 324  
 Db 301 ETSKGVYCLQYDDOKIYVSGLDNNTIKIMDKNTLECKRILGHTGVSVCLOYDERVITITSS 360  
 QY 325 SDSYIRVWDVNTGEMLNTLIHCEAVLHLRFNNGMVTGSKDRIAVWDMASPTDITLRR 384  
 Db 361 SDSYIRVWDVNTGEMLNTLIHCEAVLHLRFNNGMVTGSKDRIAVWDMASPTDITLRR 420  
 QY 385 VLVGHRAAVNVVDPDDKTYVASGDRITIKVWNTSTCEPRTLNGHRGIAICLOYDRILV 444  
 Db 421 VLVGHRAAVNVVDPDDKTYVASGDRITIKVWNTSTCEPRTLNGHRGIAICLOYDRILV 480

QY 445 SSSDNTIRLMDIEGACIRLVGHEELVRCIRPDNKRIVSGADGKIKVMDVLAALDPR 504  
 |||||||  
 Db 481 SSSDNTIRLMDIEGACIRLVGHEELVRCIRPDNKRIVSGADGKIKVMDVLAALDPR 540  
 QY 505 AAGTCLCTIVLHSGVRVRLQDFEFOIVSSSHDPTILIMFLNDPAAQAEPSPSRKTY 564  
 |||||||  
 Db 541 AAGTCLCTIVLHSGVRVRLQDFEFOIVSSSHDPTILIMFLNDPAAQAEPSPSRKTY 600  
 QY 565 TYISR 569  
 |||||  
 Db 601 TYISR 605

RESULT 4  
 US-09-724-676-52729  
 ; Sequence 52729, Application US/09724676  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676  
 ; CURRENT FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 97222  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO: 52729  
 ; LENGTH: 537  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-724-676-52729

Query Match 89.0%; Score 2701; DB 5; Length 537;

Best Local Similarity 91.2%; Pred. No. 3.8e-242;

Matches 519; Conservative 4; Mismatches 14; Indels 32; Gaps 2;

QY 1 MDPAAVLOEKALKFNSSEREDCNNGEPKRIPEKNSLRQTYNSCARCLINOETVCLA 60  
 |||||||  
 Db 1 MDPAAVLOEKALKFNSSEREDCNNGEPKRIPEKNSLRQTYNSCARCLINOETVCLA 60  
 QY 61 STAKTENCVAKTKLANGTSMIVPKORKLSAYEKEKELCVYFEQWESDQVEFEVHL 120  
 |||||||  
 Db 61 STAKTENCVAKTKLANGTSMIVPKORKLSAYEKEKELCVYFEQWESDQVEFEVHL 120  
 QY 121 ISOMCHYOHGHINSYKPKMLQDFITLAPARGLDHIAENITLSDAKSLCAAEVCKEY 180  
 |||||||  
 Db 121 ISOMCHYOHGHINSYKPKMLQDFITLAPARGLDHIAENITLSDAKSLCAAEVCKEY 180  
 QY 181 RYSDGMIMKLLERAVRTDLSIMGLAERGWGQYLEFKNRPDGNAPNSFYRATPKII 240  
 |||||||  
 Db 181 RYSDGMIMKLLERAVRTDLSIMGLAERGWGQYLEFKNRPDGNAPNSFYRATPKII 240  
 QY 241 ODIEETIESNMRCGRHSIORICRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTLECK 300  
 |||||||  
 Db 241 ODIEETIESNMRCGRHSIORICRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTLECK 300  
 QY 213 ---TIESNMRCGRHSIORICRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTLECK 268  
 |||||||  
 Db 213 ---TIESNMRCGRHSIORICRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTLECK 268  
 QY 301 RILTGHTGSVLCLOYDERVIITGSSDSTVRYVDVNTGEMLTLIHCEAVLHLRFNNGM 360  
 |||||||  
 Db 301 RILTGHTGSVLCLOYDERVIITGSSDSTVRYVDVNTGEMLTLIHCEAVLHLRFNNGM 360  
 QY 269 RILTGHTGSVLCLOYDERVIITGSSDSTVRYVDVNTGEMLTLIHCEAVLHLRFNNGM 328  
 |||||||  
 Db 269 RILTGHTGSVLCLOYDERVIITGSSDSTVRYVDVNTGEMLTLIHCEAVLHLRFNNGM 328  
 QY 361 VTCSDRSIAVWDASPTDITLRLVGVHRAAVNVVDDDKYIYASAGDRTIKVWNTSTC 420  
 |||||||  
 Db 361 VTCSDRSIAVWDASPTDITLRLVGVHRAAVNVVDDDKYIYASAGDRTIKVWNTSTC 420  
 QY 329 VTCSDRSIAVWDASPTDITLRLVGVHRAAVNVVDDDKYIYASAGDRTIKVWNTSTC 388  
 |||||||  
 Db 329 VTCSDRSIAVWDASPTDITLRLVGVHRAAVNVVDDDKYIYASAGDRTIKVWNTSTC 388  
 QY 421 EFVRTLNGHKRGIACTLOYRDRLVVGSSDNTIRLMDIEGACIRLVGHEELVRCIRPDN 480  
 |||||||  
 Db 421 EFVRTLNGHKRGIACTLOYRDRLVVGSSDNTIRLMDIEGACIRLVGHEELVRCIRPDN 480  
 QY 389 EFVRTLNGHKRGIACTLOYRDRLVVGSSDNTIRLMDIEGACIRLVGHEELVRCIRPDN 448  
 |||||||  
 Db 389 EFVRTLNGHKRGIACTLOYRDRLVVGSSDNTIRLMDIEGACIRLVGHEELVRCIRPDN 448  
 QY 481 KRIVSAYGKIKVMDVLAALDPRAPAGTCLRTLVHSGVRVRLQDFEFOIVSSSHDT 540  
 |||||||  
 Db 481 KRIVSAYGKIKVMDVLAALDPRAPAGTCLRTLVHSGVRVRLQDFEFOIVSSSHDT 540  
 QY 449 KRIVSAYGKIKVMDVLAALDPRAPAGTCLRTLVHSGVRVRLQDFEFOIVSSSHDT 508  
 |||||||  
 Db 449 KRIVSAYGKIKVMDVLAALDPRAPAGTCLRTLVHSGVRVRLQDFEFOIVSSSHDT 508  
 QY 541 ILIMFLNDPAAQAEPSPSRKTYISR 569  
 |||||||  
 Db 541 ILIMFLNDPAAQAEPSPSRKTYISR 569  
 QY 509 ILIMFLNDPAAQAEPSPSRKTYISR 537

RESULT 5  
 US-09-724-676-52730  
 ; Sequence 52730, Application US/09724676  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Compugen LTD  
 ; TITLE OF INVENTION: Variants of alternative splicing  
 ; FILE REFERENCE: 129181.4 Compugen  
 ; CURRENT APPLICATION NUMBER: US/09/724,676  
 ; CURRENT FILING DATE: 2000-11-28  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO: 52730  
 ; LENGTH: 537  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-724-676-52730

Query Match 89.0%; Score 2701; DB 5; Length 537;

Best Local Similarity 91.2%; Pred. No. 3.8e-242;

Matches 519; Conservative 4; Mismatches 14; Indels 32; Gaps 2;

QY 1 MDPAAVLOEKALKFNSSEREDCNNGEPKRIPEKNSLRQTYNSCARCLINOETVCLA 60  
 |||||||  
 Db 1 MDPAAVLOEKALKFNSSEREDCNNGEPKRIPEKNSLRQTYNSCARCLINOETVCLA 60  
 QY 61 STAKTENCVAKTKLANGTSMIVPKORKLSAYEKEKELCVYFEQWESDQVEFEVHL 120  
 |||||||  
 Db 61 STAKTENCVAKTKLANGTSMIVPKORKLSAYEKEKELCVYFEQWESDQVEFEVHL 120  
 QY 121 ISOMCHYOHGHINSYKPKMLQDFITLAPARGLDHIAENITLSDAKSLCAAEVCKEY 180  
 |||||||  
 Db 121 ISOMCHYOHGHINSYKPKMLQDFITLAPARGLDHIAENITLSDAKSLCAAEVCKEY 180  
 QY 181 RYSDGMIMKLLERAVRTDLSIMGLAERGWGQYLEFKNRPDGNAPNSFYRATPKII 240  
 |||||||  
 Db 181 RYSDGMIMKLLERAVRTDLSIMGLAERGWGQYLEFKNRPDGNAPNSFYRATPKII 240  
 QY 241 ODIEETIESNMRCGRHSIORICRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTLECK 300  
 |||||||  
 Db 241 ODIEETIESNMRCGRHSIORICRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTLECK 300  
 QY 213 ---TIESNMRCGRHSIORICRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTLECK 268  
 |||||||  
 Db 213 ---TIESNMRCGRHSIORICRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTLECK 268  
 QY 301 RILTGHTGSVLCLOYDERVIITGSSDSTVRYVDVNTGEMLTLIHCEAVLHLRFNNGM 360  
 |||||||  
 Db 301 RILTGHTGSVLCLOYDERVIITGSSDSTVRYVDVNTGEMLTLIHCEAVLHLRFNNGM 360  
 QY 269 RILTGHTGSVLCLOYDERVIITGSSDSTVRYVDVNTGEMLTLIHCEAVLHLRFNNGM 328  
 |||||||  
 Db 269 RILTGHTGSVLCLOYDERVIITGSSDSTVRYVDVNTGEMLTLIHCEAVLHLRFNNGM 328  
 QY 361 VTCSDRSIAVWDASPTDITLRLVGVHRAAVNVVDDDKYIYASAGDRTIKVWNTSTC 420  
 |||||||  
 Db 361 VTCSDRSIAVWDASPTDITLRLVGVHRAAVNVVDDDKYIYASAGDRTIKVWNTSTC 420  
 QY 329 VTCSDRSIAVWDASPTDITLRLVGVHRAAVNVVDDDKYIYASAGDRTIKVWNTSTC 388  
 |||||||  
 Db 329 VTCSDRSIAVWDASPTDITLRLVGVHRAAVNVVDDDKYIYASAGDRTIKVWNTSTC 388  
 QY 421 EFVRTLNGHKRGIACTLOYRDRLVVGSSDNTIRLMDIEGACIRLVGHEELVRCIRPDN 480  
 |||||||  
 Db 421 EFVRTLNGHKRGIACTLOYRDRLVVGSSDNTIRLMDIEGACIRLVGHEELVRCIRPDN 480  
 QY 389 EFVRTLNGHKRGIACTLOYRDRLVVGSSDNTIRLMDIEGACIRLVGHEELVRCIRPDN 448  
 |||||||  
 Db 389 EFVRTLNGHKRGIACTLOYRDRLVVGSSDNTIRLMDIEGACIRLVGHEELVRCIRPDN 448  
 QY 481 KRIVSAYGKIKVMDVLAALDPRAPAGTCLRTLVHSGVRVRLQDFEFOIVSSSHDT 540  
 |||||||  
 Db 481 KRIVSAYGKIKVMDVLAALDPRAPAGTCLRTLVHSGVRVRLQDFEFOIVSSSHDT 540  
 QY 449 KRIVSAYGKIKVMDVLAALDPRAPAGTCLRTLVHSGVRVRLQDFEFOIVSSSHDT 508  
 |||||||  
 Db 449 KRIVSAYGKIKVMDVLAALDPRAPAGTCLRTLVHSGVRVRLQDFEFOIVSSSHDT 508  
 QY 541 ILIMFLNDPAAQAEPSPSRKTYISR 569  
 |||||||  
 Db 541 ILIMFLNDPAAQAEPSPSRKTYISR 569  
 QY 509 ILIMFLNDPAAQAEPSPSRKTYISR 537

RESULT 6

US-09-724-676A-52729

; Sequence 52729, Application US/09724676A

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676A

; CURRENT FILING DATE: 2000-11-28

NUMBER OF SEQ ID NOS: 97222  
 SOFTWARE: Patentin version 3.2  
 SEQ ID NO 52729  
 LENGTH: 537  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-724-676A-52730

Query Match 89.0%; Score 2701; DB 5; Length 537;  
 Best Local Similarity 91.2%; Pred. No. 3.8e-242;  
 Matches 519; Conservative 4; Mismatches 14; Indels 32; Gaps 2;

```

QY 1 MDPAAVLOEKALFKMNSEREDCNGEPPRKIIPEKNSLRQYNSCARCLINQETVCLA 60
DB 1 MDPAAVLOEKALFKMNSEREDCNGEPPRKIIPEKNSLRQYNSCARCLINQETVCLA 60
QY 61 STAMKTEHCYAKTKRLANGSSMIVPQKRLSASYEKEKELCVKFEQWSESDOVEFEVHL 120
DB 61 STAMKTEHCYAKTKRLANGSSMIVPQKRLSASYEKEKELCVKFEQWSESDOVEFEVHL 120
QY 121 ISOMCHYOHGHTINSYKPMLOQDFITLALPARGLDHIAENILSYDAKSLCAAEIVCKEMY 180
DB 121 ISOMCHYOHGHTINSYKPMLOQDFITLALPARGLDHIAENILSYDAKSLCAAEIVCKEMY 180
QY 181 RVTSQDMLMKKLIERNVTRDSLMRGLAERGGCYLFKNKPPDGNAPNSFYRALYPRKI 240
DB 181 RVTSQDMLMKKLIERNVTRDSLMRGLAERGGCYLFKNKPPDGNAPNSFYRALYPRKI 240
QY 241 ODLETESNMRCGRHSLOIRHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
DB 241 ODLETESNMRCGRHSLOIRHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
QY 213 ---TTESNMRCGRHSLOIRHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECK 268
DB 213 ---TTESNMRCGRHSLOIRHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECK 268
QY 301 RLTGHTGSVLCLOYDERVITIGSSDSTVRYWVDVTGEMLNTLIIHCEAVLHLRFNNGM 360
DB 301 RLTGHTGSVLCLOYDERVITIGSSDSTVRYWVDVTGEMLNTLIIHCEAVLHLRFNNGM 360
QY 269 RLTGHTGSVLCLOYDERVITIGSSDSTVRYWVDVTGEMLNTLIIHCEAVLHLRFNNGM 328
DB 269 RLTGHTGSVLCLOYDERVITIGSSDSTVRYWVDVTGEMLNTLIIHCEAVLHLRFNNGM 328
QY 361 VTCSDKRSIAVWDMASPDITLRLVGVHRAAVNVVDFDXYIVASGDRITIKWNTSTC 420
DB 361 VTCSDKRSIAVWDMASPDITLRLVGVHRAAVNVVDFDXYIVASGDRITIKWNTSTC 420
QY 329 VTCSDKRSIAVWDMASPDITLRLVGVHRAAVNVVDFDXYIVASGDRITIKWNTSTC 388
DB 329 VTCSDKRSIAVWDMASPDITLRLVGVHRAAVNVVDFDXYIVASGDRITIKWNTSTC 388
QY 421 EFWRLNGHKKRIACLOYRDLRVVSGSSDNTIRLMDIECGACLVLEGHELYRCIRFDN 480
DB 421 EFWRLNGHKKRIACLOYRDLRVVSGSSDNTIRLMDIECGACLVLEGHELYRCIRFDN 480
QY 481 KRIYGAVDGKIKWDVLAALDPRAPAGTLCFLTVESHGVRFLQDFEQIVSSSHDT 540
DB 481 KRIYGAVDGKIKWDVLAALDPRAPAGTLCFLTVESHGVRFLQDFEQIVSSSHDT 540
QY 541 ILIWFELNDPAAQAEPSPSRITTYISR 569
DB 541 ILIWFELNDPAAQAEPSPSRITTYISR 569
QY 509 ILIWFELNDPAAQAEPSPSRITTYISR 537
DB 509 ILIWFELNDPAAQAEPSPSRITTYISR 537

```

## RESULT 7

US-09-724-676A-52730  
 Sequence 52730, Application US/09724676A  
 GENERAL INFORMATION:  
 APPLICANT: Compugen LTD  
 TITLE OF INVENTION: Variants of alternative splicing  
 FILE REFERENCE: 129181.4 Compugen  
 CURRENT APPLICATION NUMBER: US/09/724,676A  
 NUMBER OF SEQ ID NOS: 97222  
 SOFTWARE: Patentin version 3.2  
 SEQ ID NO 52730  
 LENGTH: 537  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-724-676A-52730

Query Match 89.0%; Score 2701; DB 5; Length 537;  
 Best Local Similarity 91.2%; Pred. No. 3.8e-242;  
 Matches 519; Conservative 4; Mismatches 14; Indels 32; Gaps 2;

```

QY 1 MDPAAVLOEKALFKMNSEREDCNGEPPRKIIPEKNSLRQYNSCARCLINQETVCLA 60
DB 1 MDPAAVLOEKALFKMNSEREDCNGEPPRKIIPEKNSLRQYNSCARCLINQETVCLA 60
QY 61 STAMKTEHCYAKTKRLANGSSMIVPQKRLSASYEKEKELCVKFEQWSESDOVEFEVHL 120
DB 61 STAMKTEHCYAKTKRLANGSSMIVPQKRLSASYEKEKELCVKFEQWSESDOVEFEVHL 120
QY 121 ISOMCHYOHGHTINSYKPMLOQDFITLALPARGLDHIAENILSYDAKSLCAAEIVCKEMY 180
DB 121 ISOMCHYOHGHTINSYKPMLOQDFITLALPARGLDHIAENILSYDAKSLCAAEIVCKEMY 180
QY 181 RVTSQDMLMKKLIERNVTRDSLMRGLAERGGCYLFKNKPPDGNAPNSFYRALYPRKI 240
DB 181 RVTSQDMLMKKLIERNVTRDSLMRGLAERGGCYLFKNKPPDGNAPNSFYRALYPRKI 240
QY 241 ODLETESNMRCGRHSLOIRHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
DB 241 ODLETESNMRCGRHSLOIRHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
QY 213 ---TTESNMRCGRHSLOIRHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECK 268
DB 213 ---TTESNMRCGRHSLOIRHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECK 268
QY 301 RLTGHTGSVLCLOYDERVITIGSSDSTVRYWVDVTGEMLNTLIIHCEAVLHLRFNNGM 360
DB 301 RLTGHTGSVLCLOYDERVITIGSSDSTVRYWVDVTGEMLNTLIIHCEAVLHLRFNNGM 360
QY 269 RLTGHTGSVLCLOYDERVITIGSSDSTVRYWVDVTGEMLNTLIIHCEAVLHLRFNNGM 328
DB 269 RLTGHTGSVLCLOYDERVITIGSSDSTVRYWVDVTGEMLNTLIIHCEAVLHLRFNNGM 328
QY 361 VTCSDKRSIAVWDMASPDITLRLVGVHRAAVNVVDFDXYIVASGDRITIKWNTSTC 420
DB 361 VTCSDKRSIAVWDMASPDITLRLVGVHRAAVNVVDFDXYIVASGDRITIKWNTSTC 420
QY 329 VTCSDKRSIAVWDMASPDITLRLVGVHRAAVNVVDFDXYIVASGDRITIKWNTSTC 388
DB 329 VTCSDKRSIAVWDMASPDITLRLVGVHRAAVNVVDFDXYIVASGDRITIKWNTSTC 388
QY 421 EFWRLNGHKKRIACLOYRDLRVVSGSSDNTIRLMDIECGACLVLEGHELYRCIRFDN 480
DB 421 EFWRLNGHKKRIACLOYRDLRVVSGSSDNTIRLMDIECGACLVLEGHELYRCIRFDN 480
QY 481 KRIYGAVDGKIKWDVLAALDPRAPAGTLCFLTVESHGVRFLQDFEQIVSSSHDT 540
DB 481 KRIYGAVDGKIKWDVLAALDPRAPAGTLCFLTVESHGVRFLQDFEQIVSSSHDT 540
QY 541 ILIWFELNDPAAQAEPSPSRITTYISR 569
DB 541 ILIWFELNDPAAQAEPSPSRITTYISR 569
QY 509 ILIWFELNDPAAQAEPSPSRITTYISR 537
DB 509 ILIWFELNDPAAQAEPSPSRITTYISR 537

```

## RESULT 8

US-09-724-676-52732  
 Sequence 52732, Application US/09724676  
 GENERAL INFORMATION:  
 APPLICANT: Compugen LTD  
 TITLE OF INVENTION: Variants of alternative splicing  
 FILE REFERENCE: 129181.4 Compugen  
 CURRENT APPLICATION NUMBER: US/09/724,676  
 NUMBER OF SEQ ID NOS: 97222  
 SOFTWARE: Patentin version 3.2  
 SEQ ID NO 52732  
 LENGTH: 573  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-724-676-52732

Query Match 88.1%; Score 2673; DB 5; Length 573;  
 Best Local Similarity 85.8%; Pred. No. 1.7e-239;  
 Matches 519; Conservative 4; Mismatches 14; Indels 68; Gaps 3;

```

QY 1 MDPAAVLOEKALFKMNSEREDCNGEPPRKIIPEKNSLRQYNSCARCLINQETVCLA 60
DB 1 MDPAAVLOEKALFKMNSEREDCNGEPPRKIIPEKNSLRQYNSCARCLINQETVCLA 60
QY 25 NNEPPEKRIIPEKNSLRQYNSCARCLINQETVCLASTAMKTEHCYAKTKLANGTSSMIV 84
DB 61 NNEPPEKRIIPEKNSLRQYNSCARCLINQETVCLASTAMKTEHCYAKTKLANGTSSMIV 120
QY 85 PKORKLSASYEKEKELCVKFEQWSESDOVEFEVHLISOMCHYOHGHTINSYKPMLOQDF 144
DB 85 PKORKLSASYEKEKELCVKFEQWSESDOVEFEVHLISOMCHYOHGHTINSYKPMLOQDF 144

```

```
Db 121 PKORKLSASYEKELCVYFEQWSESDOVEVEHLISOMCHYQHGHINSYLKPMLOQDF 180
Qy 145 ITALPARGLDHIENILSYLDAKSLCAELVCKEYRVTSDDGMKLIERNVRTDSLNR 204
Db 181 ITALPARGLDHIENILSYLDAKSLCAELVCKEYRVTSDDGMKLIERNVRTDSLNR 236
Qy 205 GLAERGMGOYLFRKKNPPDGNAPNSFYRALYPKIIQDIETTESMWRGGRHSIORIHRS 264
Db 237 FSVBERGRTKRM-----TIESMWRGGRHSIORIHRS 268
Qy 265 ETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECKRIITGHTGSLVCLQYDERVITGS 324
Db 269 ETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECKRIITGHTGSLVCLQYDERVITGS 328
Qy 325 SDSTVRVMDVNTGEMNLTIHHCBAVLHFRNNGMAYTCSKDRSLAVMDMSPDITLRR 384
Db 329 SDSTVRVMDVNTGEMNLTIHHCBAVLHFRNNGMAYTCSKDRSLAVMDMSPDITLRR 388
Qy 385 VLVGHRAAVNVVDFDRTKYVSASGDRITKWNVTSTCEVFTLNGHKGRIACLOYRDLRV 444
Db 389 VLVGHRAAVNVVDFDRTKYVSASGDRITKWNVTSTCEVFTLNGHKGRIACLOYRDLRV 448
Qy 445 SSSSDNTIRLMDIEGACLRVLEGEHEELVRCIRPDNKRIVSGAVDGIKVMDLVAALDPR 504
Db 449 SSSSDNTIRLMDIEGACLRVLEGEHEELVRCIRPDNKRIVSGAVDGIKVMDLVAALDPR 508
Qy 505 APAGTLCRLTIVESHSGRVFLQDFEFOIVSSSHDDTILIMDFLNDPAAQAEPPRSPRTY 564
Db 509 APAGTLCRLTIVESHSGRVFLQDFEFOIVSSSHDDTILIMDFLNDPAAQAEPPRSPRTY 568
Qy 565 TYISR 569
Db 569 TYISR 573

RESULT 9
US-09-724-676-52733
; Sequence 52733, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 52733
; LENGTH: 573
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-52733

Query Match      88.1%; Score 2673; DB 5; Length 573;
Best Local Similarity 85.8%; Pred. No. 1.7e-239;
Matches 519; Conservative 4; Mismatches 14; Indels 68; Gaps 3;

Qy 1 MDPAAVLOEKALKFM-----NSSEREDC 24
Db 1 MDPAAVLOEKALKFMCSMPRSLWLGSSSLADSMPSLRCLYNPGTALTAFOSSSEREDC 60
Qy 25 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTENCVAATKLANGTSSMIV 84
Db 61 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTENCVAATKLANGTSSMIV 120
Qy 85 PKORKLSASYEKELCVYFEQWSESDOVEVEHLISOMCHYQHGHINSYLKPMLOQDF 144
Db 121 PKORKLSASYEKELCVYFEQWSESDOVEVEHLISOMCHYQHGHINSYLKPMLOQDF 180
Qy 145 ITALPARGLDHIENILSYLDAKSLCAELVCKEYRVTSDDGMKLIERNVRTDSLNR 204
Db 181 ITALPARGLDHIENILSYLDAKSLCAELVCKEYRVTSDDGMKLIERNVRTDSLNR 236
Qy 205 GLAERGMGOYLFRKKNPPDGNAPNSFYRALYPKIIQDIETTESMWRGGRHSIORIHRS 264
```

```
Db 237 FSVBERGRTKRM-----TIESMWRGGRHSIORIHRS 268
Qy 265 ETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECKRIITGHTGSLVCLQYDERVITGS 324
Db 269 ETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECKRIITGHTGSLVCLQYDERVITGS 328
Qy 325 SDSTVRVMDVNTGEMNLTIHHCBAVLHFRNNGMAYTCSKDRSLAVMDMSPDITLRR 384
Db 329 SDSTVRVMDVNTGEMNLTIHHCBAVLHFRNNGMAYTCSKDRSLAVMDMSPDITLRR 388
Qy 385 VLVGHRAAVNVVDFDRTKYVSASGDRITKWNVTSTCEVFTLNGHKGRIACLOYRDLRV 444
Db 389 VLVGHRAAVNVVDFDRTKYVSASGDRITKWNVTSTCEVFTLNGHKGRIACLOYRDLRV 448
Qy 445 SSSSDNTIRLMDIEGACLRVLEGEHEELVRCIRPDNKRIVSGAVDGIKVMDLVAALDPR 504
Db 449 SSSSDNTIRLMDIEGACLRVLEGEHEELVRCIRPDNKRIVSGAVDGIKVMDLVAALDPR 508
Qy 505 APAGTLCRLTIVESHSGRVFLQDFEFOIVSSSHDDTILIMDFLNDPAAQAEPPRSPRTY 564
Db 509 APAGTLCRLTIVESHSGRVFLQDFEFOIVSSSHDDTILIMDFLNDPAAQAEPPRSPRTY 568
Qy 565 TYISR 569
Db 569 TYISR 573

RESULT 10
US-09-724-676A-52732
; Sequence 52732, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 52732
; LENGTH: 573
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-52732

Query Match      88.1%; Score 2673; DB 5; Length 573;
Best Local Similarity 85.8%; Pred. No. 1.7e-239;
Matches 519; Conservative 4; Mismatches 14; Indels 68; Gaps 3;

Qy 1 MDPAAVLOEKALKFM-----NSSEREDC 24
Db 1 MDPAAVLOEKALKFMCSMPRSLWLGSSSLADSMPSLRCLYNPGTALTAFOSSSEREDC 60
Qy 25 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTENCVAATKLANGTSSMIV 84
Db 61 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTENCVAATKLANGTSSMIV 120
Qy 85 PKORKLSASYEKELCVYFEQWSESDOVEVEHLISOMCHYQHGHINSYLKPMLOQDF 144
Db 121 PKORKLSASYEKELCVYFEQWSESDOVEVEHLISOMCHYQHGHINSYLKPMLOQDF 180
Qy 145 ITALPARGLDHIENILSYLDAKSLCAELVCKEYRVTSDDGMKLIERNVRTDSLNR 204
Db 181 ITALPARGLDHIENILSYLDAKSLCAELVCKEYRVTSDDGMKLIERNVRTDSLNR 236
Qy 205 GLAERGMGOYLFRKKNPPDGNAPNSFYRALYPKIIQDIETTESMWRGGRHSIORIHRS 264
Db 237 FSVBERGRTKRM-----TIESMWRGGRHSIORIHRS 268
Qy 265 ETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECKRIITGHTGSLVCLQYDERVITGS 324
Db 269 ETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECKRIITGHTGSLVCLQYDERVITGS 328
```

QY	325	SDSIYRVMDVNTGEMLNTLIHHCEAVLHLRFNNQMMNTYCSKDRSTAYWDMA	SPDITLRR	384
Db	329	SDSIYRVMDVNTGEMLNTLIHHCEAVLHLRFNNQMMNTYCSKDRSLAYWDMA	SPDITLRR	368
QY	385	VLVGHRAAVNVNVDDEDKITYASASDRTIKWNTSTCFEFTVLNKHKGIACTLOY	RDRLVV	444
Db	389	VLVGHRAAVNVNVDDEDKITYASASDRTIKWNTSTCFEFTVLNKHKGIACTLOY	RDRLVV	448
QY	445	SGSSDNTIRLMDIECGACLRVLEHHEELVRCIRFDNKRIVSGADGKIKWDL	VLAALDPR	504
Db	449	SGSSDNTIRLMDIECGACLRVLEHHEELVRCIRFDNKRIVSGADGKIKWDL	VLAALDPR	508
QY	505	APAGTLCRLTLEHSGRVRLQDPDEFIVSSSHDITLIMDFLNDPAAQAEP	PRPSRTY	564
Db	509	APAGTLCRLTLEHSGRVRLQDPDEFIVSSSHDITLIMDFLNDPAAQAEP	PRPSRTY	568
QY	565	TYISR	569	
Db	569	TYISR	573	

RESULT 11  
US-09-724-676A-52733  
; Sequence 52733, Application US/09724676A  
; Chemical: Tetracycline

Query Match	88.1%;	Score 2673;	DB 5;	Length 573;
Best Local Similarity	85.8%;	Pred. No. 1.7e-239;		
Matches 519; Conservative	4;	Mismatches 14;	Indels 68;	Gaps 3

QY	445	SGSSSNTIRLMDICGACLRVLSEHEELVRICIRFDNKRIRVSGADGKIRKWDVLAALDPR	504
Db	449	SGSSSNTIRLMDIEGACLRVLSEHEELVRICIRFDNKRIRVSGADGKIRKWDVLAALDPR	508
QY	505	APAGTLCIRTLYEHSGRVFRLOPPEFOLVSSHDDTLIMDFLNDPAAQAEPPSPSRTY	564
Db	509	APAGTLCIRTLYEHSGRVFRLOPPEFOLVSSHDDTLIMDFLNDPAAQAEPPSPSRTY	568
QY	565	TYISR 569	
Db	569	TYISR 573	

RESULT 12  
US-09-724-676-61015  
; Sequence 61015, Application US/09724676  
; GENERAL INFORMATION.

```

; TYPE: PRI
; ORGANISM: Homo sapiens
US-09-724-676-61015

```

Db 501 ILIWDPLNPPSAQNETRSPSRTYYISR 529

RESULT 13

US-09-724-61016  
; Sequence 61016, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724, 676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 61016  
; LENGTH: 529  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-61016

Query Match 79.1%; Score 2400; DB 5; Length 529;

Best Local Similarity 79.3%; Pred. No. 3.3e-214;  
Matches 451; Conservative 43; Mismatches 35; Indels 40; Gaps 4;

Qy 1 MDPAEAVLOEALKFANSSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNOETVCLA 60  
Db 1 MEP-DSVIEDKTIELMNTSVMEDQNEDESPK-----KNTLMQ----- 36  
Qy 61 STAMKENCVAKTKLANGSSMIYVPOKRTLSAYEKEKELCVKFEQMSDOVEFEHL 120  
Db 37 -----ISNGTSSVIVSRKRPSEGNQOKERKDLCTIKFYDQWSESDQVEFEHL 82  
Qy 121 ISOMCHYOHGHINSYKLPMLQRFITLALPARGLDHAENILSYLDAKSLCAAEVCKEMY 180  
Db 83 ISMCHYOHGHINSYKLPMLQRFITLALPARGLDHAENILSYLDAKSLCAAEVCKEMQ 142  
Qy 181 RVTSDGMLMKLIERVTRDLSLRGLAERRGQGYLFKNKPPDGNAPPSNYFALYPKII 240  
Db 143 RVISGMLMKLIERVTRDPLMKGLSERRGWDGYLFKNRPDGG--PPNSFYRSILYPKII 200  
Qy 241 QDIETTESNMRGRHSIORHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300  
Db 201 QDIETTESNMRGRHNIQRIQCRSENSKGYVCLQYDDOKIISGLRDNISIKIMDKNTLECK 260  
Qy 301 RIITGHTGSVLCLOYDERVITIGSSDSTVYVWDVNTGEMLNTLIHCEAVLHLRFNNGM 360  
Db 261 KVLGHTGSVLCLOYDERVITIGSSDSTVYVWDVNTGEMLNTLIHCEAVLHLRFNNGM 320  
Qy 361 VTCSKDRSTIAVWDMASPTDITLRVLYGHRAAVNVDFDDKITYSASGDRITKWNSTJC 420  
Db 321 VTCSKDRSTIAVWDMASPTDITLRVLYGHRAAVNVDFDDKITYSASGDRITKWNSTJC 380  
Qy 421 EFWRTLNGHKGRIACLOYDRRLVYVSSDNTIRLMDIEGACRLVEGHEELVRCIRFDN 480  
Db 381 EFWRTLNGHKGRIACLOYDRRLVYVSSDNTIRLMDIEGACRLVEGHEELVRCIRFDN 440  
Qy 481 KRIYSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVEHSGRVFRLQDFEQIVSSSHDT 540  
Db 441 KRIYSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVEHSGRVFRLQDFEQIVSSSHDT 500  
Qy 541 ILIWDPLNDPAQAQEPSPSRTYYISR 569  
Db 501 ILIWDPLNPPSAQNETRSPSRTYYISR 529

RESULT 14

US-09-724-676A-61015  
; Sequence 61015, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724, 676A

; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 61015  
; LENGTH: 529  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-61015

Query Match 79.1%; Score 2400; DB 5; Length 529;

Best Local Similarity 79.3%; Pred. No. 3.3e-214;  
Matches 451; Conservative 43; Mismatches 35; Indels 40; Gaps 4;

Qy 1 MDPAEAVLOEALKFANSSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNOETVCLA 60  
Db 1 MEP-DSVIEDKTIELMNTSVMEDQNEDESPK-----KNTLMQ----- 36  
Qy 61 STAMKENCVAKTKLANGSSMIYVPOKRTLSAYEKEKELCVKFEQMSDOVEFEHL 120  
Db 37 -----ISNGTSSVIVSRKRPSEGNQOKERKDLCTIKFYDQWSESDQVEFEHL 82  
Qy 121 ISOMCHYOHGHINSYKLPMLQRFITLALPARGLDHAENILSYLDAKSLCAAEVCKEMY 180  
Db 83 ISMCHYOHGHINSYKLPMLQRFITLALPARGLDHAENILSYLDAKSLCAAEVCKEMQ 142  
Qy 181 RVTSDGMLMKLIERVTRDLSLRGLAERRGQGYLFKNKPPDGNAPPSNYFALYPKII 240  
Db 143 RVISGMLMKLIERVTRDPLMKGLSERRGWDGYLFKNRPDGG--PPNSFYRSILYPKII 200  
Qy 241 QDIETTESNMRGRHSIORHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300  
Db 201 QDIETTESNMRGRHNIQRIQCRSENSKGYVCLQYDDOKIISGLRDNISIKIMDKNTLECK 260  
Qy 301 RIITGHTGSVLCLOYDERVITIGSSDSTVYVWDVNTGEMLNTLIHCEAVLHLRFNNGM 360  
Db 261 KVLGHTGSVLCLOYDERVITIGSSDSTVYVWDVNTGEMLNTLIHCEAVLHLRFNNGM 320  
Qy 361 VTCSKDRSTIAVWDMASPTDITLRVLYGHRAAVNVDFDDKITYSASGDRITKWNSTJC 420  
Db 321 VTCSKDRSTIAVWDMASPTDITLRVLYGHRAAVNVDFDDKITYSASGDRITKWNSTJC 380  
Qy 421 EFWRTLNGHKGRIACLOYDRRLVYVSSDNTIRLMDIEGACRLVEGHEELVRCIRFDN 480  
Db 381 EFWRTLNGHKGRIACLOYDRRLVYVSSDNTIRLMDIEGACRLVEGHEELVRCIRFDN 440  
Qy 481 KRIYSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVEHSGRVFRLQDFEQIVSSSHDT 540  
Db 441 KRIYSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVEHSGRVFRLQDFEQIVSSSHDT 500  
Qy 541 ILIWDPLNDPAQAQEPSPSRTYYISR 569  
Db 501 ILIWDPLNPPSAQNETRSPSRTYYISR 529

RESULT 15

US-09-724-676A-61016  
; Sequence 61016, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724, 676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 61016  
; LENGTH: 529  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-61016

Query Match 79.1%; Score 2400; DB 5; Length 529;

Best Local Similarity 79.3%; Pred. No. 3.3e-214;

Matches 451; Conservative 43; Mismatches 35; Indels 40; Gaps 4;

```
QY 1 MDPAAVLQEKALKFANSSERECNNGEPPKIIPEKNSLRQTYNSCARLQNETVYCLA 60
Db 1 MEP-DSVLEDKTTELMTSVWEDONEDSPK-----KNTLMQ----- 36
QY 61 STAKTENCYAKTKLANGTSSMIYPKOKRLSAYEKEKELCVYFEQWSESDQVEFEYHL 120
Db 37 -----ISNGTSSVIVSRKRPESEGNQKEKDLCKYFDQWSESDQVEFEYHL 82
QY 121 ISQCHYOHGHINSYKLPMLQRPDFTALPARGLDHAENTILSYLDAKSLCAAEIVCKEWY 180
Db 83 ISRMCHYOHGHINSYKLPMLQRPDFTALPEGLDHAENTILSYLDAKSLCAAEIVCKEWQ 142
QY 181 RYISDGMIMKKLIERMVRTDSIMRGIAERGNQOYLFKNKPPDGNAPPNSTYALYPKII 240
Db 143 RYISDGMIMKKLIERMVRTDSIMRGIAERGNQOYLFKNKPPDGNAPPNSTYALYPKII 200
QY 241 QDIETIESNMRCGRHSIQRIHCHSETSKGYCQYDDQKIVSGLRDNTIKIMDKNTLECK 300
Db 201 QDIETIESNMRCGRHNQRIQCRSENGKYCQYDDQKIIISGLRDNSTIKIMDKNTLECK 260
QY 301 RILFGHTGSVLCIQYDERVITITGSSDSTVAVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
Db 261 KVLFGHTGSVLCIQYDERVITITGSSDSTVAVMDVNTGEVLNLTLIHNEAVLHLRFNNGLM 320
QY 361 VTCSKDRSIAMVMDASPTDITLRRLVGHRAAVNVYDFDDKYIVSASGDRITKVMNTSTC 420
Db 321 VTCSKDRSIAMVMDASPTDITLRRLVGHRAAVNVYDFDDKYIVSASGDRITKVMNTSTC 380
QY 421 EFWRTLNGHRRGIACIQYRDRLVYSSSDNTIRLMDIECGACLRVLEGHHELVRCIRFDN 480
Db 381 EFWRTLNGHRRGIACIQYRDRLVYSSSDNTIRLMDIECGACLRVLEGHHELVRCIRFDN 440
QY 481 KRIYSGAYDGIKIVMDLVAAALDPRAPAGTLCITLVEHSGRVPRLODFEQIVSSSHDT 540
Db 441 KRIYSGAYDGIKIVMDLVAAALDPRAPASTLCITLVEHSGRVPRLODFEQIISSHDT 500
QY 541 ILIMDFLNDPAQAQEPSPRTTYTISR 569
Db 501 ILIMDFLNDPAQAQEPSPRTTYTISR 529
```

Search completed: February 20, 2003, 09:59:48  
Job time : 59 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 09:55:51 ; Search time 149 Seconds  
(Without alignments)  
2462.102 Million cells updates/sec

Title: US-09-601-168b-2  
Perfect score: 3034  
Sequence: 1 MDPAAVLQERALKRKNSSSE.....PAAQAEPRPSKRTTYISR 569

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents AA.Main:\*  
1: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3034	100.0	569	1	PCT-US02-00311-2
2	3034	100.0	569	1	PCT-US99-19560-2
3	3034	100.0	569	16	US-09-210-060-18
4	3034	100.0	569	17	US-09-385-219-2
5	3034	100.0	569	18	US-09-415-795-4
6	3034	100.0	569	18	US-09-455-371-2

7	3034	100.0	569	18	US-09-455-371b-2	Sequence 2, Appli
8	3034	100.0	569	20	US-09-601-168b-2	Sequence 2, Appli
9	3034	100.0	569	21	US-09-700-444-12	Sequence 12, Appl
10	3034	100.0	569	21	US-09-791-537-80169	Sequence 80169, A
11	3034	100.0	569	22	US-09-832-161-181	Sequence 18, Appl
12	3034	100.0	569	24	US-10-038-010-8	Sequence 8, Appli
13	3034	100.0	569	24	US-10-042-417-2	Sequence 2, Appli
14	3034	100.0	569	27	US-60-098-1010-5	Sequence 5, Appli
15	3034	100.0	569	27	US-60-098-353-2	Sequence 2, Appli
16	3006	99.1	605	1	PCT-US01-04098A-1244	Sequence 1244, Ap
17	3006	99.1	605	21	US-09-791-537-122825	Sequence 12, Ap
18	2997	98.8	569	18	US-09-415-795-12	Sequence 12, Appl
19	2997	98.8	569	21	US-09-791-537-131810	Sequence 131810, A
20	2992.5	98.6	632	1	PCT-US01-04098A-1246	Sequence 1246, Ap
21	2990	98.5	569	21	US-09-791-537-21214	Sequence 21214, A
22	2989.5	98.5	654	1	PCT-US01-04098A-3212	Sequence 3212, Ap
23	2989.5	98.5	654	1	PCT-US01-04098A-3213	Sequence 3213, Ap
24	2989.5	98.5	654	1	PCT-US01-04098A-3214	Sequence 3214, Ap
25	2978	98.2	569	21	US-09-791-537-129637	Sequence 129637, A
26	2882	95.0	579	1	PCT-US01-04098A-1245	Sequence 1245, Ap
27	2597	85.6	518	21	US-09-791-537-83876	Sequence 83876, A
28	2597	85.6	518	27	US-60-098-355-26	Sequence 26, Appl
29	2582.5	85.1	517	1	PCT-US95-01210-30	Sequence 30, Appl
30	2400	79.1	529	21	US-09-791-537-36840	Sequence 36840, A
31	2384.5	78.6	542	1	PCT-US01-04098A-1789	Sequence 1789, Ap
32	2384.5	78.6	542	16	US-09-210-060-16	Sequence 16, Appl
33	2384.5	78.6	542	18	US-09-488-725A-3353	Sequence 3353, Ap
34	2384.5	78.6	542	21	US-09-791-537-11197	Sequence 11197, A
35	2384.5	78.6	542	22	US-09-832-161-16	Sequence 16, Appl
36	2384.5	78.6	550	1	PCT-US01-08656-9623	Sequence 9623, Ap
37	2384.5	78.6	550	18	US-09-488-725A-6925	Sequence 6925, Ap
38	2384.5	78.6	550	21	US-09-791-537-25894	Sequence 25894, A
39	2372.5	78.2	550	1	PCT-US01-14827-12977	Sequence 12977, A
40	2358	77.7	508	21	US-09-791-537-36821	Sequence 36821, A
41	2089.5	68.9	510	20	US-09-614-150-6363	Sequence 6363, Ap
42	2089.5	68.9	510	21	US-09-791-537-121061	Sequence 121061, A
43	2089.5	68.9	510	21	US-09-791-537-127362	Sequence 127362, A
44	2089.5	68.9	510	27	US-60-098-355-27	Sequence 27, Appl
45	2089.5	68.9	510	27	US-60-191-637-6384	Sequence 6384, Ap

ALIGNMENTS

RESULT 1  
PCT-US02-00311-2  
Sequence 2, Application PC/TUS0200311  
GENERAL INFORMATION:  
APPLICANT: Pagano, M.  
TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF  
TITLE OF INVENTION: PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS  
FILE REFERENCE: 5914-090-228  
CURRENT APPLICATION NUMBER: PCT/US02/00311  
CURRENT FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: 60/260,179  
PRIOR FILING DATE: 2001-01-05  
NUMBER OF SEQ ID NOS: 89  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 569  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US02-00311-2

Query Match	100.0%	Score 3034	DB 1	Length 569
Best Local Similarity	100.0%	Pred. No. 1.3e-284		
Matches	569	Conservative	0	Mismatches 0
			Indels	0
			Gaps	0
QY	1	MDPAEVLQERALKRKNSSEREDCNNGEPRIIEKNSLRQYNSCARCLINQETVCLIA	60	
DB	1	MDPAEVLQERALKRKNSSEREDCNNGEPRIIEKNSLRQYNSCARCLINQETVCLIA	60	
QY	61	STAKMTENCVAATKLANGSSMTIVPRKRLASYSKEKELCVKIFQWSESDQVEFVHL	120	

Db 61 STAKTEKCAKTLANGTSSMIVPQKRLASAEKEKELCVKFFEDWSSDVEFEHL 120  
QY 121 ISOMCHYOHGHINSYKLPMLQDRFITALPARGLDHIAENILSYLDAKSLCAELVCKEMY 180  
Db 121 ISOMCHYOHGHINSYKLPMLQDRFITALPARGLDHIAENILSYLDAKSLCAELVCKEMY 180  
QY 181 RYSDGMLMKKLIERNVTRDSLMRGLAERBGWGYLFKNKPPDGNAPNSFYRALYPKII 240  
Db 181 RYSDGMLMKKLIERNVTRDSLMRGLAERBGWGYLFKNKPPDGNAPNSFYRALYPKII 240  
QY 241 ODITTESNMRCGRHSLOHICRSETSKGYVCLQYDDOKIYSGLDNMTIKIMDKNTLECK 300  
Db 241 ODITTESNMRCGRHSLOHICRSETSKGYVCLQYDDOKIYSGLDNMTIKIMDKNTLECK 300  
QY 301 RILTGHTSVLCLQYDERVITIGSSDSTVRWVDVTGEMLNTLIHCEAVLHLRFNNGM 360  
Db 301 RILTGHTSVLCLQYDERVITIGSSDSTVRWVDVTGEMLNTLIHCEAVLHLRFNNGM 360  
QY 361 VTCSKDRSLAVWDMASPTDITLRLVGLHRAAVNVVDDKRYIASASGDRITKWNSTC 420  
Db 361 VTCSKDRSLAVWDMASPTDITLRLVGLHRAAVNVVDDKRYIASASGDRITKWNSTC 420  
QY 421 EFWRLNGHKGKIGIACLOYDRDLVYSGSSDNTIRLMDIEGACLRVLEGEHELVCIRFDN 480  
Db 421 EFWRLNGHKGKIGIACLOYDRDLVYSGSSDNTIRLMDIEGACLRVLEGEHELVCIRFDN 480  
QY 481 KRIYSGAYDGKIKWMDIYVLAALDPRAPAGTICLRTLVHSGRVRLQDFEFOIVSSSHDT 540  
Db 481 KRIYSGAYDGKIKWMDIYVLAALDPRAPAGTICLRTLVHSGRVRLQDFEFOIVSSSHDT 540  
QY 541 ILIWDPLNDPAQAQEPSPSRRTTYISR 569  
Db 541 ILIWDPLNDPAQAQEPSPSRRTTYISR 569

RESULT 2  
PCT-US99-19560-2  
Sequence 2, Application PC/TUS9919560  
GENERAL INFORMATION:  
APPLICANT: NEW YORK UNIVERSITY  
TITLE OF INVENTION: NOVEL DBIQUITTIN LIGASES AS THERAPEUTIC TARGETS  
FILE REFERENCE: 5914-081-228  
CURRENT APPLICATION NUMBER: PCT/US99/19560  
EARLIER FILING DATE: 1999-08-31  
EARLIER APPLICATION NUMBER: 60/098,355  
EARLIER FILING DATE: 1998-08-28  
EARLIER APPLICATION NUMBER: 60/118,568  
EARLIER FILING DATE: 1999-02-03  
EARLIER APPLICATION NUMBER: 60/124,449  
EARLIER FILING DATE: 1999-03-15  
NUMBER OF SEQ ID NOS: 89  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 2  
LENGTH: 569  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US99-19560-2

Query Match 100.0%; Score 3034; DB 1; Length 569;  
Best Local Similarity 100.0%; Pred. No. 1.3e-284;  
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAAVLOERKALKFNNSSEREDCNNGEPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA 60  
Db 1 MDPAAVLOERKALKFNNSSEREDCNNGEPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA 60  
QY 61 STAKTEKCAKTLKANGTSSMIVPQKRLASAEKEKELCVKFFEDWSSDVEFEHL 120  
Db 61 STAKTEKCAKTLKANGTSSMIVPQKRLASAEKEKELCVKFFEDWSSDVEFEHL 120  
QY 121 ISOMCHYOHGHINSYKLPMLQDRFITALPARGLDHIAENILSYLDAKSLCAELVCKEMY 180  
Db 121 ISOMCHYOHGHINSYKLPMLQDRFITALPARGLDHIAENILSYLDAKSLCAELVCKEMY 180

Db 121 ISOMCHYOHGHINSYKLPMLQDRFITALPARGLDHIAENILSYLDAKSLCAELVCKEMY 180  
QY 181 RYSDGMLMKKLIERNVTRDSLMRGLAERBGWGYLFKNKPPDGNAPNSFYRALYPKII 240  
Db 181 RYSDGMLMKKLIERNVTRDSLMRGLAERBGWGYLFKNKPPDGNAPNSFYRALYPKII 240  
QY 241 ODITTESNMRCGRHSLOHICRSETSKGYVCLQYDDOKIYSGLDNMTIKIMDKNTLECK 300  
Db 241 ODITTESNMRCGRHSLOHICRSETSKGYVCLQYDDOKIYSGLDNMTIKIMDKNTLECK 300  
QY 301 RILTGHTSVLCLQYDERVITIGSSDSTVRWVDVTGEMLNTLIHCEAVLHLRFNNGM 360  
Db 301 RILTGHTSVLCLQYDERVITIGSSDSTVRWVDVTGEMLNTLIHCEAVLHLRFNNGM 360  
QY 361 VTCSKDRSLAVWDMASPTDITLRLVGLHRAAVNVVDDKRYIASASGDRITKWNSTC 420  
Db 361 VTCSKDRSLAVWDMASPTDITLRLVGLHRAAVNVVDDKRYIASASGDRITKWNSTC 420  
QY 421 EFWRLNGHKGKIGIACLOYDRDLVYSGSSDNTIRLMDIEGACLRVLEGEHELVCIRFDN 480  
Db 421 EFWRLNGHKGKIGIACLOYDRDLVYSGSSDNTIRLMDIEGACLRVLEGEHELVCIRFDN 480  
QY 481 KRIYSGAYDGKIKWMDIYVLAALDPRAPAGTICLRTLVHSGRVRLQDFEFOIVSSSHDT 540  
Db 481 KRIYSGAYDGKIKWMDIYVLAALDPRAPAGTICLRTLVHSGRVRLQDFEFOIVSSSHDT 540  
QY 541 ILIWDPLNDPAQAQEPSPSRRTTYISR 569  
Db 541 ILIWDPLNDPAQAQEPSPSRRTTYISR 569

RESULT 3  
US-09-210-060-18  
Sequence 18, Application US/09210060  
GENERAL INFORMATION:  
APPLICANT: Manning, Anthony M.  
APPLICANT: Mercurio, Frank  
APPLICANT: Amlt, Sharon  
APPLICANT: Ben-Neriah, Yimon  
APPLICANT: Davis, Matti  
APPLICANT: Hatzubai, Ada  
APPLICANT: Lanyon, Iris  
APPLICANT: Yaron, Avraham  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING ACTIVATION OF  
FILE REFERENCE: NF-KB  
CURRENT APPLICATION NUMBER: US/09/210,060  
EARLIER FILING DATE: 1998-12-10  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 18  
LENGTH: 569  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-210-060-18

Query Match 100.0%; Score 3034; DB 16; Length 569;  
Best Local Similarity 100.0%; Pred. No. 1.3e-284;  
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAAVLOERKALKFNNSSEREDCNNGEPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA 60  
Db 1 MDPAAVLOERKALKFNNSSEREDCNNGEPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA 60  
QY 61 STAKTEKCAKTLKANGTSSMIVPQKRLASAEKEKELCVKFFEDWSSDVEFEHL 120  
Db 61 STAKTEKCAKTLKANGTSSMIVPQKRLASAEKEKELCVKFFEDWSSDVEFEHL 120  
QY 121 ISOMCHYOHGHINSYKLPMLQDRFITALPARGLDHIAENILSYLDAKSLCAELVCKEMY 180  
Db 121 ISOMCHYOHGHINSYKLPMLQDRFITALPARGLDHIAENILSYLDAKSLCAELVCKEMY 180  
QY 181 RYSDGMLMKKLIERNVTRDSLMRGLAERBGWGYLFKNKPPDGNAPNSFYRALYPKII 240  
Db 181 RYSDGMLMKKLIERNVTRDSLMRGLAERBGWGYLFKNKPPDGNAPNSFYRALYPKII 240

```

Db 181 RYVSDGMLKKLIERVVRTDLSLGRGLAERRGQYLFKKRPPDGNAPNSFYRALYPKII 240
QY 241 ODIEETIESNMRCGRHSIORHCRSETSKGVYCLQYDDOKIVSGLRDNITKIMDNLTLECK 300
Db 241 ODIEETIESNMRCGRHSIORHCRSETSKGVYCLQYDDOKIVSGLRDNITKIMDNLTLECK 300
QY 301 RLITGHTGVSVCLOQYDERVIITGSSDSTVRWVDVNTGEMLNTLHHCBAVHLRFNNGM 360
Db 301 RLITGHTGVSVCLOQYDERVIITGSSDSTVRWVDVNTGEMLNTLHHCBAVHLRFNNGM 360
QY 361 VTCSDKRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDDKIVYASAGDRITKVMNTSTC 420
Db 361 VTCSDKRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDDKIVYASAGDRITKVMNTSTC 420
QY 421 EFPRTLNHKKRGIACTQYDRILVYVSSSDNTIRLMDIEGACLRVLEGEHELVCIRFDN 480
Db 421 EFPRTLNHKKRGIACTQYDRILVYVSSSDNTIRLMDIEGACLRVLEGEHELVCIRFDN 480
QY 481 KRIVSGAYDGKIKVMDLVAALDPRAPACTLCLRTLVHSGRVFRLQDFEFOIVSSSHDT 540
Db 481 KRIVSGAYDGKIKVMDLVAALDPRAPACTLCLRTLVHSGRVFRLQDFEFOIVSSSHDT 540
QY 541 ILIMDFLNDPAAQAEPSPRSRTTYISR 569
Db 541 ILIMDFLNDPAAQAEPSPRSRTTYISR 569

```

RESULT 4

```

US-09-385-219-2
; Sequence 2, Application US/09385219
; GENERAL INFORMATION:
; APPLICANT: Chlaur, D.
; APPLICANT: Pagano, M.
; APPLICANT: Laires, E.
; TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS
; FILE REFERENCE: 5914-081
; CURRENT APPLICATION NUMBER: US/09/385,219
; EARLIER FILING DATE: 1999-08-27
; EARLIER APPLICATION NUMBER: 60/098,355
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: 60/118,568
; EARLIER FILING DATE: 1999-02-03
; EARLIER APPLICATION NUMBER: 60/124,449
; EARLIER FILING DATE: 1999-03-15
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-385-219-2

```

Query Match 100.0%; Score 3034; DB 17; Length 569;

Best Local Similarity 100.0%; Pred. No. 1.3e-284; Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MDPBAVLOEKALKFNNSSEREDCNNGEPRIIPKNSLRQTYNSCARLCLNOETVCLA 60
Db 1 MDPBAVLOEKALKFNNSSEREDCNNGEPRIIPKNSLRQTYNSCARLCLNOETVCLA 60
QY 61 STAMKTENCVAATKTLANGTSSMIVPKOKRLASVYKEKEKLVYFEQMSSEDQVEFEHL 120
Db 61 STAMKTENCVAATKTLANGTSSMIVPKOKRLASVYKEKEKLVYFEQMSSEDQVEFEHL 120
QY 121 ISOMCHYOHGHINSYKPMLODFITLALPARGLDHAENILSYLDAKSLCAAEIVCKEMY 180
Db 121 ISOMCHYOHGHINSYKPMLODFITLALPARGLDHAENILSYLDAKSLCAAEIVCKEMY 180
QY 181 RYVSDGMLKKLIERVVRTDLSLGRGLAERRGQYLFKKRPPDGNAPNSFYRALYPKII 240
Db 181 RYVSDGMLKKLIERVVRTDLSLGRGLAERRGQYLFKKRPPDGNAPNSFYRALYPKII 240

```

```

QY 241 ODIEETIESNMRCGRHSIORHCRSETSKGVYCLQYDDOKIVSGLRDNITKIMDNLTLECK 300
Db 241 ODIEETIESNMRCGRHSIORHCRSETSKGVYCLQYDDOKIVSGLRDNITKIMDNLTLECK 300
QY 301 RLITGHTGVSVCLOQYDERVIITGSSDSTVRWVDVNTGEMLNTLHHCBAVHLRFNNGM 360
Db 301 RLITGHTGVSVCLOQYDERVIITGSSDSTVRWVDVNTGEMLNTLHHCBAVHLRFNNGM 360
QY 361 VTCSDKRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDDKIVYASAGDRITKVMNTSTC 420
Db 361 VTCSDKRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDDKIVYASAGDRITKVMNTSTC 420
QY 421 EFPRTLNHKKRGIACTQYDRILVYVSSSDNTIRLMDIEGACLRVLEGEHELVCIRFDN 480
Db 421 EFPRTLNHKKRGIACTQYDRILVYVSSSDNTIRLMDIEGACLRVLEGEHELVCIRFDN 480
QY 481 KRIVSGAYDGKIKVMDLVAALDPRAPACTLCLRTLVHSGRVFRLQDFEFOIVSSSHDT 540
Db 481 KRIVSGAYDGKIKVMDLVAALDPRAPACTLCLRTLVHSGRVFRLQDFEFOIVSSSHDT 540
QY 541 ILIMDFLNDPAAQAEPSPRSRTTYISR 569
Db 541 ILIMDFLNDPAAQAEPSPRSRTTYISR 569

```

RESULT 5

```

US-09-415-795-4
; Sequence 4, Application US/09415795
; GENERAL INFORMATION:
; APPLICANT: Zhou, Pengbo
; APPLICANT: Howley, Peter M.
; TITLE OF INVENTION: TARGETED PROTEOLYSIS BY RECRUITMENT TO UBIQUITIN
; TITLE OF INVENTION: PROTEIN LIGASES
; FILE REFERENCE: HMV-043.01
; CURRENT APPLICATION NUMBER: US/09/415,795
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-415-795-4

```

Query Match 100.0%; Score 3034; DB 18; Length 569;

Best Local Similarity 100.0%; Pred. No. 1.3e-284; Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MDPBAVLOEKALKFNNSSEREDCNNGEPRIIPKNSLRQTYNSCARLCLNOETVCLA 60
Db 1 MDPBAVLOEKALKFNNSSEREDCNNGEPRIIPKNSLRQTYNSCARLCLNOETVCLA 60
QY 61 STAMKTENCVAATKTLANGTSSMIVPKOKRLASVYKEKEKLVYFEQMSSEDQVEFEHL 120
Db 61 STAMKTENCVAATKTLANGTSSMIVPKOKRLASVYKEKEKLVYFEQMSSEDQVEFEHL 120
QY 121 ISOMCHYOHGHINSYKPMLODFITLALPARGLDHAENILSYLDAKSLCAAEIVCKEMY 180
Db 121 ISOMCHYOHGHINSYKPMLODFITLALPARGLDHAENILSYLDAKSLCAAEIVCKEMY 180
QY 181 RYVSDGMLKKLIERVVRTDLSLGRGLAERRGQYLFKKRPPDGNAPNSFYRALYPKII 240
Db 181 RYVSDGMLKKLIERVVRTDLSLGRGLAERRGQYLFKKRPPDGNAPNSFYRALYPKII 240
QY 241 ODIEETIESNMRCGRHSIORHCRSETSKGVYCLQYDDOKIVSGLRDNITKIMDNLTLECK 300
Db 241 ODIEETIESNMRCGRHSIORHCRSETSKGVYCLQYDDOKIVSGLRDNITKIMDNLTLECK 300
QY 301 RLITGHTGVSVCLOQYDERVIITGSSDSTVRWVDVNTGEMLNTLHHCBAVHLRFNNGM 360
Db 301 RLITGHTGVSVCLOQYDERVIITGSSDSTVRWVDVNTGEMLNTLHHCBAVHLRFNNGM 360
QY 361 VTCSDKRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDDKIVYASAGDRITKVMNTSTC 420

```

```

Db 361 VTGSKDRIAWMDASPDITLRLVGVGHRAAVNVDFDKYIVASGDRITKWNSTC 420
QY 421 EFVRLNGHKKRGIACTQYDRLVVSGSSDNTIRLMDIEGACRLVLEGHHELVRCIRPDN 480
Db 421 EFVRLNGHKKRGIACTQYDRLVVSGSSDNTIRLMDIEGACRLVLEGHHELVRCIRPDN 480
QY 481 KRIVSGAYDGKIKWMDLVVALDPRAPAGTLCRTLVESHGVRFLQDFEFOIVSSHDDT 540
Db 481 KRIVSGAYDGKIKWMDLVVALDPRAPAGTLCRTLVESHGVRFLQDFEFOIVSSHDDT 540
QY 541 ILIWDPLNDPAQAQAEPPRSPSRRTTYISR 569
Db 541 ILIWDPLNDPAQAQAEPPRSPSRRTTYISR 569

```

# RESULT 6

```

US-09-455-371-2
; Sequence 2, Application US/09455371
; GENERAL INFORMATION:
; APPLICANT: Beer-Romero, Peggy
; APPLICANT: Ellledge, Stephen
; APPLICANT: Harper, J. Wade
; APPLICANT: Rolfe, Mark
; APPLICANT: Strack, Peter J.
; APPLICANT: Winston, Jeffrey T.
; TITLE OF INVENTION: Regulation of I Kappa B (Ik-B) Degradation and Methods and Reagents
; FILE REFERENCE: 120541-1010
; CURRENT APPLICATION NUMBER: US/09/455,371
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/110982
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent version 3.1
; SEQ ID NO 2
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-455-371-2

```

Query Match 100.0%; Score 3034; DB 18; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-284;  
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MDPAEAVLQEKALKFMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNDETVCIA 60
Db 1 MDPAEAVLQEKALKFMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNDETVCIA 60
QY 61 STAMKTENCVAKTILANGTSSMIVPKORKLSASYEKEKELCVKFFQWSESDQVEFEHL 120
Db 61 STAMKTENCVAKTILANGTSSMIVPKORKLSASYEKEKELCVKFFQWSESDQVEFEHL 120
QY 121 ISOMCHYOHGHINSYLRKPMQRODFTTALPARGLDHAENTLSTYDAKSLCAAEIVCKEM 180
Db 121 ISOMCHYOHGHINSYLRKPMQRODFTTALPARGLDHAENTLSTYDAKSLCAAEIVCKEM 180
QY 181 RYVSDGMLMKKLIEMRVRTDSLMRGLAERGMGQYLFKNKPPDGNAPNSFYRALYPRII 240
Db 181 RYVSDGMLMKKLIEMRVRTDSLMRGLAERGMGQYLFKNKPPDGNAPNSFYRALYPRII 240
QY 241 QDIETTESNMRCGRHSIORIHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
Db 241 QDIETTESNMRCGRHSIORIHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
QY 301 RILTGHTGSVLCLQYDERVITIGSSDSTVYRWVDVTGEMLNTLIHCEAVLHLRFNNGM 360
Db 301 RILTGHTGSVLCLQYDERVITIGSSDSTVYRWVDVTGEMLNTLIHCEAVLHLRFNNGM 360
QY 361 VTGSKDRIAWMDASPDITLRLVGVGHRAAVNVDFDKYIVASGDRITKWNSTC 420
Db 361 VTGSKDRIAWMDASPDITLRLVGVGHRAAVNVDFDKYIVASGDRITKWNSTC 420

```

```

QY 421 EFVRLNGHKKRGIACTQYDRLVVSGSSDNTIRLMDIEGACRLVLEGHHELVRCIRPDN 480
Db 421 EFVRLNGHKKRGIACTQYDRLVVSGSSDNTIRLMDIEGACRLVLEGHHELVRCIRPDN 480
QY 481 KRIVSGAYDGKIKWMDLVVALDPRAPAGTLCRTLVESHGVRFLQDFEFOIVSSHDDT 540
Db 481 KRIVSGAYDGKIKWMDLVVALDPRAPAGTLCRTLVESHGVRFLQDFEFOIVSSHDDT 540
QY 541 ILIWDPLNDPAQAQAEPPRSPSRRTTYISR 569
Db 541 ILIWDPLNDPAQAQAEPPRSPSRRTTYISR 569

```

# RESULT 7

```

US-09-455-371B-2
; Sequence 2, Application US/09455371B
; GENERAL INFORMATION:
; APPLICANT: Beer-Romero, Peggy
; APPLICANT: Ellledge, Stephen
; APPLICANT: Harper, J. Wade
; APPLICANT: Rolfe, Mark
; APPLICANT: Strack, Peter J.
; APPLICANT: Winston, Jeffrey T.
; TITLE OF INVENTION: Regulation of I Kappa B (Ik-B) Degradation and Methods and Reagents
; FILE REFERENCE: 120541-1010
; CURRENT APPLICATION NUMBER: US/09/455,371B
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/110982
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent version 3.1
; SEQ ID NO 2
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-455-371B-2

```

Query Match 100.0%; Score 3034; DB 18; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-284;  
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MDPAEAVLQEKALKFMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNDETVCIA 60
Db 1 MDPAEAVLQEKALKFMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNDETVCIA 60
QY 61 STAMKTENCVAKTILANGTSSMIVPKORKLSASYEKEKELCVKFFQWSESDQVEFEHL 120
Db 61 STAMKTENCVAKTILANGTSSMIVPKORKLSASYEKEKELCVKFFQWSESDQVEFEHL 120
QY 121 ISOMCHYOHGHINSYLRKPMQRODFTTALPARGLDHAENTLSTYDAKSLCAAEIVCKEM 180
Db 121 ISOMCHYOHGHINSYLRKPMQRODFTTALPARGLDHAENTLSTYDAKSLCAAEIVCKEM 180
QY 181 RYVSDGMLMKKLIEMRVRTDSLMRGLAERGMGQYLFKNKPPDGNAPNSFYRALYPRII 240
Db 181 RYVSDGMLMKKLIEMRVRTDSLMRGLAERGMGQYLFKNKPPDGNAPNSFYRALYPRII 240
QY 241 QDIETTESNMRCGRHSIORIHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
Db 241 QDIETTESNMRCGRHSIORIHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300
QY 301 RILTGHTGSVLCLQYDERVITIGSSDSTVYRWVDVTGEMLNTLIHCEAVLHLRFNNGM 360
Db 301 RILTGHTGSVLCLQYDERVITIGSSDSTVYRWVDVTGEMLNTLIHCEAVLHLRFNNGM 360
QY 361 VTGSKDRIAWMDASPDITLRLVGVGHRAAVNVDFDKYIVASGDRITKWNSTC 420
Db 361 VTGSKDRIAWMDASPDITLRLVGVGHRAAVNVDFDKYIVASGDRITKWNSTC 420
QY 421 EFVRLNGHKKRGIACTQYDRLVVSGSSDNTIRLMDIEGACRLVLEGHHELVRCIRPDN 480
Db 421 EFVRLNGHKKRGIACTQYDRLVVSGSSDNTIRLMDIEGACRLVLEGHHELVRCIRPDN 480

```

QY 481 KRIVSGAYGKIKWDLVAALDPRAPAGTLCRTLYVHSGVFRLODEFOIVSSHDDT 540  
DB 481 KRIVSGAYGKIKWDLVAALDPRAPAGTLCRTLYVHSGVFRLODEFOIVSSHDDT 540  
QY 541 ILIMDFLNDPAAOAEPPRSPRTTYTISR 569  
DB 541 ILIMDFLNDPAAOAEPPRSPRTTYTISR 569

## RESULT 8

US-09-601-168a-2  
; Sequence 2, Application US/09601168A  
; GENERAL INFORMATION:  
; APPLICANT: BENAROUS, Richard  
; APPLICANT: MARGOTTIN, Florence  
; APPLICANT: DURAND, Hervé  
; APPLICANT: ARENZANA SEISDEDOS, Fernando  
; APPLICANT: KROLL, Mathias  
; APPLICANT: CONDORET, Jean-Paul  
; TITLE OF INVENTION: Human beta-TTCP protein  
; FILE REFERENCE: 935,38812X00  
; CURRENT APPLICATION NUMBER: US/09/601,168A  
; CURRENT FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: PCT/FR99/00196  
; PRIOR FILING DATE: 1999-01-29  
; PRIOR APPLICATION NUMBER: FR98 01100  
; PRIOR FILING DATE: 1998-01-30  
; PRIOR APPLICATION NUMBER: FR98 15545  
; PRIOR FILING DATE: 1998-12-09  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 569  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Description of the artificial sequence : CDNA  
; OTHER INFORMATION: coding for human beta-TTCP protein  
US-09-601-168a-2

Query Match 100.0%; Score 3034; DB 20; Length 569;  
Best Local Similarity 100.0%; Pred. No. 1.3e-284;  
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDPBAVLOEKALKFMNSEREDCNNGEPPRKIIPEKNSLRQTYNSCARCLINQETVCLA 60  
DB 1 MDPBAVLOEKALKFMNSEREDCNNGEPPRKIIPEKNSLRQTYNSCARCLINQETVCLA 60  
QY 61 STAKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKFFQWSESDQVEFEVHL 120  
DB 61 STAKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKFFQWSESDQVEFEVHL 120  
QY 121 ISOMCHYOHGHIINSYLPKMLORDFTLPAAGLDHIAENLSTLIDAKSLCAAEVLCYKEMV 180  
DB 121 ISOMCHYOHGHIINSYLPKMLORDFTLPAAGLDHIAENLSTLIDAKSLCAAEVLCYKEMV 180  
QY 181 RYTSOGMLMKKLIBRMVYTDLSLWGLAERRGWCYLFKNRPPDGNAPPNSFYBALYPKII 240  
DB 181 RYTSOGMLMKKLIBRMVYTDLSLWGLAERRGWCYLFKNRPPDGNAPPNSFYBALYPKII 240  
QY 241 ODIEFTESNMRCGRHSIORHICRSETSKGYVCLQYDDOKIVSGIRDMTIIKIMDKNTLECK 300  
DB 241 ODIEFTESNMRCGRHSIORHICRSETSKGYVCLQYDDOKIVSGIRDMTIIKIMDKNTLECK 300  
QY 301 RLUTGHTGSVLCLOQDEVEVIITGSSDSIVRYVDVNTGEMLNTLIHHCBAVLHLRFNNGM 360  
DB 301 RLUTGHTGSVLCLOQDEVEVIITGSSDSIVRYVDVNTGEMLNTLIHHCBAVLHLRFNNGM 360  
QY 361 VTCSKDRSIAVWDMASPTDITLRVLYGHRAAVNVDFDKYIVSASGDRTIKVMNTSTC 420  
DB 361 VTCSKDRSIAVWDMASPTDITLRVLYGHRAAVNVDFDKYIVSASGDRTIKVMNTSTC 420

QY 421 EFVRLNGHKGRIAGLOQYRDLVYSSGSDNTIRLMDIEGACALRYLEGHELYNCIRFDN 480  
DB 421 EFVRLNGHKGRIAGLOQYRDLVYSSGSDNTIRLMDIEGACALRYLEGHELYNCIRFDN 480  
QY 481 KRIVSGAYGKIKWDLVAALDPRAPAGTLCRTLYVHSGVFRLODEFOIVSSHDDT 540  
DB 481 KRIVSGAYGKIKWDLVAALDPRAPAGTLCRTLYVHSGVFRLODEFOIVSSHDDT 540  
QY 541 ILIMDFLNDPAAOAEPPRSPRTTYTISR 569  
DB 541 ILIMDFLNDPAAOAEPPRSPRTTYTISR 569

## RESULT 9

US-09-700-444-12  
; Sequence 12, Application US/09700444  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE PHARMACEUTICALS, INC.  
; APPLICANT: BANDMAN, Olga  
; APPLICANT: HILMAN, Jennifer L.  
; APPLICANT: LAL, Preeti  
; APPLICANT: YUE, Henry  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: PATTERSON, Chandra  
; APPLICANT: BAUGHN, Marian R.  
; TITLE OF INVENTION: CELL SIGNALING PROTEINS  
; FILE REFERENCE: PF-0521 PCT  
; CURRENT APPLICATION NUMBER: US/09/700,444  
; CURRENT FILING DATE: 2002-08-26  
; PRIOR APPLICATION NUMBER: 60/085,3434; 60/098,010  
; PRIOR FILING DATE: 1998-05-13; 1998-08-26  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PERL Program  
; SEQ ID NO 12  
; LENGTH: 569  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc-feature  
; OTHER INFORMATION: Incyte Clone 3239149  
US-09-700-444-12

Query Match 100.0%; Score 3034; DB 21; Length 569;  
Best Local Similarity 100.0%; Pred. No. 1.3e-284;  
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDPBAVLOEKALKFMNSEREDCNNGEPPRKIIPEKNSLRQTYNSCARCLINQETVCLA 60  
DB 1 MDPBAVLOEKALKFMNSEREDCNNGEPPRKIIPEKNSLRQTYNSCARCLINQETVCLA 60  
QY 61 STAKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKFFQWSESDQVEFEVHL 120  
DB 61 STAKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKFFQWSESDQVEFEVHL 120  
QY 121 ISOMCHYOHGHIINSYLPKMLORDFTLPAAGLDHIAENLSTLIDAKSLCAAEVLCYKEMV 180  
DB 121 ISOMCHYOHGHIINSYLPKMLORDFTLPAAGLDHIAENLSTLIDAKSLCAAEVLCYKEMV 180  
QY 181 RYTSOGMLMKKLIBRMVYTDLSLWGLAERRGWCYLFKNRPPDGNAPPNSFYBALYPKII 240  
DB 181 RYTSOGMLMKKLIBRMVYTDLSLWGLAERRGWCYLFKNRPPDGNAPPNSFYBALYPKII 240  
QY 241 ODIEFTESNMRCGRHSIORHICRSETSKGYVCLQYDDOKIVSGIRDMTIIKIMDKNTLECK 300  
DB 241 ODIEFTESNMRCGRHSIORHICRSETSKGYVCLQYDDOKIVSGIRDMTIIKIMDKNTLECK 300  
QY 301 RLUTGHTGSVLCLOQDEVEVIITGSSDSIVRYVDVNTGEMLNTLIHHCBAVLHLRFNNGM 360  
DB 301 RLUTGHTGSVLCLOQDEVEVIITGSSDSIVRYVDVNTGEMLNTLIHHCBAVLHLRFNNGM 360  
QY 361 VTCSKDRSIAVWDMASPTDITLRVLYGHRAAVNVDFDKYIVSASGDRTIKVMNTSTC 420  
DB 361 VTCSKDRSIAVWDMASPTDITLRVLYGHRAAVNVDFDKYIVSASGDRTIKVMNTSTC 420

Db 361 VTCSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKITYASGDRITKWMNTSTC 420  
 QY 421 EFVRLNKHKGACIACLODRRLVYVSSSDNTIRLMDIEGCACIRVLEGHEELVRCIRFDN 480  
 Db 421 EFVRLNKHKGACIACLODRRLVYVSSSDNTIRLMDIEGCACIRVLEGHEELVRCIRFDN 480  
 QY 481 KRIVSAGVDGKIKYWDVLAALDPRAPAGTLCRTLVHSGRVRLQDFEQIVSSSHDT 540  
 Db 481 KRIVSAGVDGKIKYWDVLAALDPRAPAGTLCRTLVHSGRVRLQDFEQIVSSSHDT 540  
 QY 541 ILIWFDLNDPAAQAEPSPRSRTTYTISR 569  
 Db 541 ILIWFDLNDPAAQAEPSPRSRTTYTISR 569

## RESULT 10

US-09-791-537-80169  
 ; Sequence 80169, Application US/09/91537  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bionomix, Inc.  
 ; APPLICANT: Debe, Derek  
 ; APPLICANT: Danzer, Joseph  
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
 ; FILE REFERENCE: 261/210  
 ; CURRENT APPLICATION NUMBER: US/09/791,537  
 ; CURRENT FILING DATE: 2001-02-22  
 ; NUMBER OF SEQ ID NOS: 153055  
 ; SOFTWARE: Patent version 3.0  
 ; SEQ ID NO 80169  
 ; LENGTH: 569  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-791-537-80169

Query Match 100.0%; Score 3034; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 1,3e-284;  
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAEAVLOEKALKFMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNDQETVCLA 60  
 Db 1 MDPAEAVLOEKALKFMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNDQETVCLA 60  
 QY 61 STAMKTEHCVAKTCLANGSSMIVPQKRLSASYEKEKELCVKFEQMSDDOYEVEHL 120  
 Db 61 STAMKTEHCVAKTCLANGSSMIVPQKRLSASYEKEKELCVKFEQMSDDOYEVEHL 120  
 QY 121 ISOMCHYOHGHINSYLKPMLOPDEFTALPARGLDHIAENILSYLDAKSLCAAEIVCKEY 180  
 Db 121 ISOMCHYOHGHINSYLKPMLOPDEFTALPARGLDHIAENILSYLDAKSLCAAEIVCKEY 180  
 QY 181 RYTSOGMLMKKLIBRMVPTDLSLWGLAERKMGQYLFKNKPPDGNAPPNFYTALPKII 240  
 Db 181 RYTSOGMLMKKLIBRMVPTDLSLWGLAERKMGQYLFKNKPPDGNAPPNFYTALPKII 240  
 QY 241 ODIEFTESNMRCGRHSIORHICRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300  
 Db 241 ODIEFTESNMRCGRHSIORHICRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300  
 QY 301 RLITGHTGSVCLQYDEVEYITGSSDSTVAVWVNTGEMLNTLIHHCBAVLHLRFNNGM 360  
 Db 301 RLITGHTGSVCLQYDEVEYITGSSDSTVAVWVNTGEMLNTLIHHCBAVLHLRFNNGM 360  
 QY 361 VTCSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKITYASGDRITKWMNTSTC 420  
 Db 361 VTCSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKITYASGDRITKWMNTSTC 420  
 QY 421 EFVRLNKHKGACIACLODRRLVYVSSSDNTIRLMDIEGCACIRVLEGHEELVRCIRFDN 480  
 Db 421 EFVRLNKHKGACIACLODRRLVYVSSSDNTIRLMDIEGCACIRVLEGHEELVRCIRFDN 480  
 QY 481 KRIVSAGVDGKIKYWDVLAALDPRAPAGTLCRTLVHSGRVRLQDFEQIVSSSHDT 540  
 Db 481 KRIVSAGVDGKIKYWDVLAALDPRAPAGTLCRTLVHSGRVRLQDFEQIVSSSHDT 540

Db 481 KRIVSAGVDGKIKYWDVLAALDPRAPAGTLCRTLVHSGRVRLQDFEQIVSSSHDT 540  
 QY 541 ILIWFDLNDPAAQAEPSPRSRTTYTISR 569  
 Db 541 ILIWFDLNDPAAQAEPSPRSRTTYTISR 569

## RESULT 11

US-09-832-161-18  
 ; Sequence 18, Application US/09832161  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Manning, Anthony M.  
 ; APPLICANT: Mercurio, Frank  
 ; APPLICANT: Amit, Sharon  
 ; APPLICANT: Ben-Neriah, Yimon  
 ; APPLICANT: Davis, Matti  
 ; APPLICANT: Hatzuba, Ada  
 ; APPLICANT: Lavon, Itis  
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING ACTIVATION OF  
 ; FILE REFERENCE: NF-KB  
 ; CURRENT APPLICATION NUMBER: US/09/832,161  
 ; CURRENT FILING DATE: 2001-04-09  
 ; PRIOR FILING DATE: 1998-12-10  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: Patent Ver. 2.0  
 ; SEQ ID NO 18  
 ; LENGTH: 569  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-832-161-18

Query Match 100.0%; Score 3034; DB 22; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 1,3e-284;  
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAEAVLOEKALKFMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNDQETVCLA 60  
 Db 1 MDPAEAVLOEKALKFMNSEREDCNNGEPKRIPEKNSLRQTYNSCARLCLNDQETVCLA 60  
 QY 61 STAMKTEHCVAKTCLANGSSMIVPQKRLSASYEKEKELCVKFEQMSDDOYEVEHL 120  
 Db 61 STAMKTEHCVAKTCLANGSSMIVPQKRLSASYEKEKELCVKFEQMSDDOYEVEHL 120  
 QY 121 ISOMCHYOHGHINSYLKPMLOPDEFTALPARGLDHIAENILSYLDAKSLCAAEIVCKEY 180  
 Db 121 ISOMCHYOHGHINSYLKPMLOPDEFTALPARGLDHIAENILSYLDAKSLCAAEIVCKEY 180  
 QY 181 RYTSOGMLMKKLIBRMVPTDLSLWGLAERKMGQYLFKNKPPDGNAPPNFYTALPKII 240  
 Db 181 RYTSOGMLMKKLIBRMVPTDLSLWGLAERKMGQYLFKNKPPDGNAPPNFYTALPKII 240  
 QY 241 ODIEFTESNMRCGRHSIORHICRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300  
 Db 241 ODIEFTESNMRCGRHSIORHICRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300  
 QY 301 RLITGHTGSVCLQYDEVEYITGSSDSTVAVWVNTGEMLNTLIHHCBAVLHLRFNNGM 360  
 Db 301 RLITGHTGSVCLQYDEVEYITGSSDSTVAVWVNTGEMLNTLIHHCBAVLHLRFNNGM 360  
 QY 361 VTCSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKITYASGDRITKWMNTSTC 420  
 Db 361 VTCSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKITYASGDRITKWMNTSTC 420  
 QY 421 EFVRLNKHKGACIACLODRRLVYVSSSDNTIRLMDIEGCACIRVLEGHEELVRCIRFDN 480  
 Db 421 EFVRLNKHKGACIACLODRRLVYVSSSDNTIRLMDIEGCACIRVLEGHEELVRCIRFDN 480  
 QY 481 KRIVSAGVDGKIKYWDVLAALDPRAPAGTLCRTLVHSGRVRLQDFEQIVSSSHDT 540  
 Db 481 KRIVSAGVDGKIKYWDVLAALDPRAPAGTLCRTLVHSGRVRLQDFEQIVSSSHDT 540

QY 541 ILIWDPLNDPAQAEPSPRSRTTYTISR 569  
Db 541 ILIWDPLNDPAQAEPSPRSRTTYTISR 569

RESULT 12

US-10-038-010-8  
; Sequence 8, Application US/10038010  
; GENERAL INFORMATION:  
; APPLICANT: HYBRIGENICS  
; APPLICANT: Pierre, legrain  
; TITLE OF INVENTION: Protein-protein interactions in adipocyte cells  
; FILE REFERENCE: B4/67A  
; CURRENT APPLICATION NUMBER: US/10/038,010  
; CURRENT FILING DATE: 2002-07-23  
; PRIOR APPLICATION NUMBER: US 60/259,377  
; PRIOR FILING DATE: 2001-01-02  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 569  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: beta-Trop  
; LOCATION: (1)..(569)  
; OTHER INFORMATION:  
US-10-038-010-8

Query Match 100.0%; Score 3034; DB 24; Length 569;  
Best Local Similarity 100.0%; Pred. No. 1.3e-284;  
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAEAVLQKALKFPMNSEREDCNNGEPKRIPEKNSLRQYNSCARLCLINQETVCLA 60  
Db 1 MDPAEAVLQKALKFPMNSEREDCNNGEPKRIPEKNSLRQYNSCARLCLINQETVCLA 60  
QY 61 STAMKTEVCVAKTKLANGTSSMIVPKOKLSASYEKEKELCVKTFEOWSESDDVEFEVHL 120  
Db 61 STAMKTEVCVAKTKLANGTSSMIVPKOKLSASYEKEKELCVKTFEOWSESDDVEFEVHL 120  
QY 121 ISOMCHYGHINSYLPKMLQDFITLALPARGLDHAENILSYLDAKSLCAAEVCKEMV 180  
Db 121 ISOMCHYGHINSYLPKMLQDFITLALPARGLDHAENILSYLDAKSLCAAEVCKEMV 180  
QY 181 RYVSDGMLMKKLIEMVRTDLSMRGLAERGGWGYLFKNKPPDGNAPNSFYRALTYPKII 240  
Db 181 RYVSDGMLMKKLIEMVRTDLSMRGLAERGGWGYLFKNKPPDGNAPNSFYRALTYPKII 240  
QY 241 QDIETTESNMRCGRHSIORHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300  
Db 241 QDIETTESNMRCGRHSIORHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300  
QY 301 RILTGHTGSVLCLODYDERVITIGSSDSYRVWDVNTGEMLNTLIHCEAVLHLRPNNGM 360  
Db 301 RILTGHTGSVLCLODYDERVITIGSSDSYRVWDVNTGEMLNTLIHCEAVLHLRPNNGM 360  
QY 361 VTCGSKDRSIYVWDMASTPTDITLRVLYGHRVAVNVDDDKYIVASGDDRTIKYVNTSTC 420  
Db 361 VTCGSKDRSIYVWDMASTPTDITLRVLYGHRVAVNVDDDKYIVASGDDRTIKYVNTSTC 420  
QY 421 EFVRLTNGHKGRIACLOYRDLRVLVSGSSDNTIRLMDIEGACLRVLEHGEELVRCIRFDN 480  
Db 421 EFVRLTNGHKGRIACLOYRDLRVLVSGSSDNTIRLMDIEGACLRVLEHGEELVRCIRFDN 480  
QY 481 KRIYSGAYDGKIKVMDVLAALDPRAPAGTLCRLTVEHSGRVFRLQDFEFOIVSSSHDT 540  
Db 481 KRIYSGAYDGKIKVMDVLAALDPRAPAGTLCRLTVEHSGRVFRLQDFEFOIVSSSHDT 540  
QY 541 ILIWDPLNDPAQAEPSPRSRTTYTISR 569  
Db 541 ILIWDPLNDPAQAEPSPRSRTTYTISR 569

RESULT 13

US-10-042-417-2  
; Sequence 2, Application US/10042417  
; GENERAL INFORMATION:  
; APPLICANT: pagano, M.  
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF  
; FILE REFERENCE: 5914-090-999  
; CURRENT APPLICATION NUMBER: US/10/042,417  
; PRIOR FILING DATE: 2002-01-07  
; NUMBER OF SEQ ID NOS: 89  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 569  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-042-417-2

Query Match 100.0%; Score 3034; DB 24; Length 569;  
Best Local Similarity 100.0%; Pred. No. 1.3e-284;  
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAEAVLQKALKFPMNSEREDCNNGEPKRIPEKNSLRQYNSCARLCLINQETVCLA 60  
Db 1 MDPAEAVLQKALKFPMNSEREDCNNGEPKRIPEKNSLRQYNSCARLCLINQETVCLA 60  
QY 61 STAMKTEVCVAKTKLANGTSSMIVPKOKLSASYEKEKELCVKTFEOWSESDDVEFEVHL 120  
Db 61 STAMKTEVCVAKTKLANGTSSMIVPKOKLSASYEKEKELCVKTFEOWSESDDVEFEVHL 120  
QY 121 ISOMCHYGHINSYLPKMLQDFITLALPARGLDHAENILSYLDAKSLCAAEVCKEMV 180  
Db 121 ISOMCHYGHINSYLPKMLQDFITLALPARGLDHAENILSYLDAKSLCAAEVCKEMV 180  
QY 181 RYVSDGMLMKKLIEMVRTDLSMRGLAERGGWGYLFKNKPPDGNAPNSFYRALTYPKII 240  
Db 181 RYVSDGMLMKKLIEMVRTDLSMRGLAERGGWGYLFKNKPPDGNAPNSFYRALTYPKII 240  
QY 241 QDIETTESNMRCGRHSIORHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300  
Db 241 QDIETTESNMRCGRHSIORHCRSETSKGYVCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300  
QY 301 RILTGHTGSVLCLODYDERVITIGSSDSYRVWDVNTGEMLNTLIHCEAVLHLRPNNGM 360  
Db 301 RILTGHTGSVLCLODYDERVITIGSSDSYRVWDVNTGEMLNTLIHCEAVLHLRPNNGM 360  
QY 361 VTCGSKDRSIYVWDMASTPTDITLRVLYGHRVAVNVDDDKYIVASGDDRTIKYVNTSTC 420  
Db 361 VTCGSKDRSIYVWDMASTPTDITLRVLYGHRVAVNVDDDKYIVASGDDRTIKYVNTSTC 420  
QY 421 EFVRLTNGHKGRIACLOYRDLRVLVSGSSDNTIRLMDIEGACLRVLEHGEELVRCIRFDN 480  
Db 421 EFVRLTNGHKGRIACLOYRDLRVLVSGSSDNTIRLMDIEGACLRVLEHGEELVRCIRFDN 480  
QY 481 KRIYSGAYDGKIKVMDVLAALDPRAPAGTLCRLTVEHSGRVFRLQDFEFOIVSSSHDT 540  
Db 481 KRIYSGAYDGKIKVMDVLAALDPRAPAGTLCRLTVEHSGRVFRLQDFEFOIVSSSHDT 540  
QY 541 ILIWDPLNDPAQAEPSPRSRTTYTISR 569  
Db 541 ILIWDPLNDPAQAEPSPRSRTTYTISR 569

RESULT 14

US-60-098-010-5  
; Sequence 5, Application US/60098010  
; GENERAL INFORMATION:  
; APPLICANT: Yue, Henry  
; APPLICANT: Tang, Y. Tom

```

; APPLICANT: Yang, Junming
; APPLICANT: Lal, Preeli
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Patterson, Chandra
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: G-PROTEIN SIGNALING PATHWAY MOLECULES
; FILE REFERENCE: PF-0583 P
; CURRENT APPLICATION NUMBER: US/60/098,010
; CURRENT FILING DATE: 1998-08-26
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 569
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 3239149
US-60-098-010-5

```

```

Query Match          100.0%; Score 3034; DB 27; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.3e-284;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MDPAEAVLQEKALKFNMSSEREDCNGEPKRIIPKNSLRQTYNSCARLCLNDETVCLA 60
DB 1 MDPAEAVLQEKALKFNMSSEREDCNGEPKRIIPKNSLRQTYNSCARLCLNDETVCLA 60
QY 61 STAKTEKNCVAKTKLANGTSSMTIVPKQKLSASYEKEKELCVKFEOMSESQVEFEVHL 120
DB 61 STAKTEKNCVAKTKLANGTSSMTIVPKQKLSASYEKEKELCVKFEOMSESQVEFEVHL 120
QY 121 ISQCHYOHGHINSYLLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVCKEMY 180
DB 121 ISQCHYOHGHINSYLLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVCKEMY 180
QY 121 ISQCHYOHGHINSYLLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVCKEMY 180
DB 121 ISQCHYOHGHINSYLLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVCKEMY 180
QY 181 RVTSDGMLMKKLLERVRVDSLMRGLAERRGQYLFKKNPPDGNAPPSFYRALYPKII 240
DB 181 RVTSDGMLMKKLLERVRVDSLMRGLAERRGQYLFKKNPPDGNAPPSFYRALYPKII 240
QY 181 RVTSDGMLMKKLLERVRVDSLMRGLAERRGQYLFKKNPPDGNAPPSFYRALYPKII 240
DB 181 RVTSDGMLMKKLLERVRVDSLMRGLAERRGQYLFKKNPPDGNAPPSFYRALYPKII 240
QY 241 ODIEETIESNMRCGRHSLORIHCRSETSKGYVCLQYDDOKIVSGLDNTIKIMDKNTLECK 300
DB 241 ODIEETIESNMRCGRHSLORIHCRSETSKGYVCLQYDDOKIVSGLDNTIKIMDKNTLECK 300
QY 301 RILTGHTGSVCLQYDERVIITGSSDSSTVVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
DB 301 RILTGHTGSVCLQYDERVIITGSSDSSTVVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
QY 361 VTCSKDRSIAVMDMASPTDITLRVLYVGHRAAVNVVDFDDKIIVSASGDRITIKWNTSTC 420
DB 361 VTCSKDRSIAVMDMASPTDITLRVLYVGHRAAVNVVDFDDKIIVSASGDRITIKWNTSTC 420
QY 421 EFVRTLNGHKGRIACLOYRDLVYVSGSSDNTIRLMDIEGACLVLEGEHELVRCIFRDN 480
DB 421 EFVRTLNGHKGRIACLOYRDLVYVSGSSDNTIRLMDIEGACLVLEGEHELVRCIFRDN 480
QY 481 KRIVSGAYDGKIKVMDLVAAIDPRAPAGTLCRLTVHSGRVRLQDFEFQIVSSSHDT 540
DB 481 KRIVSGAYDGKIKVMDLVAAIDPRAPAGTLCRLTVHSGRVRLQDFEFQIVSSSHDT 540
QY 541 ILIMDFLNDPAAQAEPPRSPSRITYIISR 569
DB 541 ILIMDFLNDPAAQAEPPRSPSRITYIISR 569

```

```

RESULT 15
US-60-098-355-2
; Sequence 2, Application US/60098355
; GENERAL INFORMATION:
; APPLICANT: Chaur, Dah Shian
; APPLICANT: Pagano, Michele
; APPLICANT: Lacher, Esther
; TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS

```

```

; FILE REFERENCE: 5914-070
; CURRENT APPLICATION NUMBER: US/60/098,355
; CURRENT FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 2
; LENGTH: 569
; TYPE: PRF
; ORGANISM: HUMAN
US-60-098-355-2

```

```

Query Match          100.0%; Score 3034; DB 27; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.3e-284;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MDPAEAVLQEKALKFNMSSEREDCNGEPKRIIPKNSLRQTYNSCARLCLNDETVCLA 60
DB 1 MDPAEAVLQEKALKFNMSSEREDCNGEPKRIIPKNSLRQTYNSCARLCLNDETVCLA 60
QY 61 STAKTEKNCVAKTKLANGTSSMTIVPKQKLSASYEKEKELCVKFEOMSESQVEFEVHL 120
DB 61 STAKTEKNCVAKTKLANGTSSMTIVPKQKLSASYEKEKELCVKFEOMSESQVEFEVHL 120
QY 121 ISQCHYOHGHINSYLLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVCKEMY 180
DB 121 ISQCHYOHGHINSYLLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVCKEMY 180
QY 121 ISQCHYOHGHINSYLLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVCKEMY 180
DB 121 ISQCHYOHGHINSYLLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVCKEMY 180
QY 181 RVTSDGMLMKKLLERVRVDSLMRGLAERRGQYLFKKNPPDGNAPPSFYRALYPKII 240
DB 181 RVTSDGMLMKKLLERVRVDSLMRGLAERRGQYLFKKNPPDGNAPPSFYRALYPKII 240
QY 181 RVTSDGMLMKKLLERVRVDSLMRGLAERRGQYLFKKNPPDGNAPPSFYRALYPKII 240
DB 181 RVTSDGMLMKKLLERVRVDSLMRGLAERRGQYLFKKNPPDGNAPPSFYRALYPKII 240
QY 241 ODIEETIESNMRCGRHSLORIHCRSETSKGYVCLQYDDOKIVSGLDNTIKIMDKNTLECK 300
DB 241 ODIEETIESNMRCGRHSLORIHCRSETSKGYVCLQYDDOKIVSGLDNTIKIMDKNTLECK 300
QY 301 RILTGHTGSVCLQYDERVIITGSSDSSTVVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
DB 301 RILTGHTGSVCLQYDERVIITGSSDSSTVVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
QY 361 VTCSKDRSIAVMDMASPTDITLRVLYVGHRAAVNVVDFDDKIIVSASGDRITIKWNTSTC 420
DB 361 VTCSKDRSIAVMDMASPTDITLRVLYVGHRAAVNVVDFDDKIIVSASGDRITIKWNTSTC 420
QY 421 EFVRTLNGHKGRIACLOYRDLVYVSGSSDNTIRLMDIEGACLVLEGEHELVRCIFRDN 480
DB 421 EFVRTLNGHKGRIACLOYRDLVYVSGSSDNTIRLMDIEGACLVLEGEHELVRCIFRDN 480
QY 481 KRIVSGAYDGKIKVMDLVAAIDPRAPAGTLCRLTVHSGRVRLQDFEFQIVSSSHDT 540
DB 481 KRIVSGAYDGKIKVMDLVAAIDPRAPAGTLCRLTVHSGRVRLQDFEFQIVSSSHDT 540
QY 541 ILIMDFLNDPAAQAEPPRSPSRITYIISR 569
DB 541 ILIMDFLNDPAAQAEPPRSPSRITYIISR 569

```

```

Search completed: February 20, 2003, 09:58:44
Job time : 151 secs

```



FT Region 467..492  
 FT /note="WD motif"  
 FT Region 516..544  
 FT /note="WD motif"  
 XX W09938969-A1.  
 XX  
 XX 05-AUG-1999.  
 XX  
 XX 29-JAN-1999; 99WC-FR00196.  
 XX  
 XX 09-DEC-1998; 98FR-0015545.  
 XX 30-JAN-1998; 98FR-0001100.  
 XX  
 XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 XX (INSP ) INST PASTEUR.  
 XX  
 XX Arenzana Seisdedos F, Benarous R, Concordet J, Durand H;  
 XX Kroll M, Margottin F;  
 XX WPI: 1999-469329/39.  
 XX N-PSDB: AAX86501.  
 XX  
 XX New human beta-transducin repeat containing protein and its  
 XX fragments useful as, or to screen for, antiviral, antitumour,  
 XX anti-inflammatory and anti-Alzheimer's agents  
 XX  
 XX Claim 1: Page 60-61; 71pp; French.  
 XX  
 XX The present sequence represents a human beta-transducin repeat containing  
 XX protein (beta-Trcp). The protein directs proteins to the proteosome  
 XX degradation pathways. The protein is able to interact with the Vpu  
 XX protein of human immune deficiency virus-1 (HIV-1), cellular proteins  
 XX IkappaB or beta-catenin (bc) and/or protein Skp1. The protein controls  
 XX ubiquitinylation of phosphorylated proteins and thus their targeting to  
 XX proteosomes for degradation. Depending on whether the process is  
 XX inhibited or promoted, the result may be delayed breakdown of CD4 (in  
 XX cases of HIV-1 infection); increased activity of Ikb (and thus reduced  
 XX activity of NFkappaB); increased degradation of mutant bc in tumour  
 XX cells, or increased bc survival (and reduced apoptosis) in Alzheimer's  
 XX patients. The beta-Trcp protein, and its active peptide fragments, or its  
 XX nucleic acid, are used to screen for anti HIV-1 agents (antivirals),  
 XX antitumour agents that disrupt cell cycle regulation or protein  
 XX degradation in human tumour cells, and anti-inflammatory agents that  
 XX disrupt activation by NFkappaB. Fragments of the protein are also  
 XX useful for treating osteo-articular inflammation or acute inflammation  
 XX associated with release of tumour necrosis factor.  
 XX  
 XX Sequence 569 AA;  
 XX  
 XX Query Match 100.0%; Score 3034; DB 20; Length 569;  
 XX Best Local Similarity 100.0%; Pred. No. 1.4e-286;  
 XX Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MDPAAVAVLQEKALFKMNSREDKNGNPPKIIPEKNSLRQVNSCARCLINQETVCLA 60  
 DB 1 MDPAAVAVLQEKALFKMNSREDKNGNPPKIIPEKNSLRQVNSCARCLINQETVCLA 60  
 QY 61 STAMTENCVAKTLANGTSSMIYVKKOKLSASTEKEKELCVKYEQWSSDQVEFEHL 120  
 DB 61 STAMTENCVAKTLANGTSSMIYVKKOKLSASTEKEKELCVKYEQWSSDQVEFEHL 120  
 QY 121 ISOMCHYHGHINSYLRQDFITLALPARGDHTIENILSTYDAKSLCAAEVYCKEWY 180  
 DB 121 ISOMCHYHGHINSYLRQDFITLALPARGDHTIENILSTYDAKSLCAAEVYCKEWY 180  
 QY 161 RYTSQGMIMKKLIERMYVTDLSLIRGLAERRGQGYLFKNRPPDGNAPPNPFYALPKII 240  
 DB 161 RYTSQGMIMKKLIERMYVTDLSLIRGLAERRGQGYLFKNRPPDGNAPPNPFYALPKII 240  
 QY 241 ODIEETESNMRCGRHSIORITCRSETSKGYCLOYDQKIVSGIRDTIKIMDKNTLECK 300  
 DB 241 ODIEETESNMRCGRHSIORITCRSETSKGYCLOYDQKIVSGIRDTIKIMDKNTLECK 300

QY 301 RILTHMGHSYVLCLOYDERVIITGSSDSTVAVWVQVNTGEMLNTLIHHCVAVLHRRNGM 360  
 DB 301 RILTHMGHSYVLCLOYDERVIITGSSDSTVAVWVQVNTGEMLNTLIHHCVAVLHRRNGM 360  
 QY 361 VTGSKDRSIAMWMASTPTDITLRVAVLGRRAAVNVYDFDKYIVASGDRITKVNSTGC 420  
 DB 361 VTGSKDRSIAMWMASTPTDITLRVAVLGRRAAVNVYDFDKYIVASGDRITKVNSTGC 420  
 QY 421 EFVRLTHGHRGTAQLOYRDLVAVSSDNTITLMDIEGACLRVLEGEELVRCITRDN 480  
 DB 421 EFVRLTHGHRGTAQLOYRDLVAVSSDNTITLMDIEGACLRVLEGEELVRCITRDN 480  
 QY 481 KRIYSGAYDGKIKVMDLVAAIDPRAPAGTLCRTLVHSGRFRLOFDFQIVSSHDT 540  
 DB 481 KRIYSGAYDGKIKVMDLVAAIDPRAPAGTLCRTLVHSGRFRLOFDFQIVSSHDT 540  
 QY 541 ILIMDFLNDPAAQAEPPRSRPTTYIISR 569  
 DB 541 ILIMDFLNDPAAQAEPPRSRPTTYIISR 569

RESULT 2  
 AAB12813  
 ID AAB12813 standard; protein; 569 AA.

AC AAB12813;  
 XX

DT 27-NOV-2000 (first entry)  
 XX

DE Human beta-transducin repeat containing protein (beta-Trcp) S60 ID NO:3.  
 XX

KM Ubiquitin ligase SCF complex; F-box protein; ubiquitinylation; IkappaB;  
 KW beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1;  
 KW gene therapy; colon cancer; beta-transducin repeat containing protein;  
 XX beta-Trcp.  
 XX

OS Homo sapiens.  
 XX

PN JP2000166542-A.  
 XX

PD 20-JUN-2000.  
 XX

PF 02-DEC-1998; 98JP-0343437.  
 XX

PR 02-DEC-1998; 98JP-0343437.  
 XX

PA (KAGA-) KAGAKU GIUTSU SHINKO JIGYODAN.  
 XX

DR WPI: 2000-485550/43.  
 XX

DR N-PSDB: AAA73132.  
 XX

PT F-box protein of ubiquitin ligase SCF complex which promotes the  
 XX ubiquitinylation of IkappaB or beta-catenin  
 XX

PS Claim 3; Page 10-12; 19pp; Japanese.  
 XX

CC The present invention describes an F-box motif protein of ubiquitin  
 CC ligase SCF complex which promotes the ubiquitinylation of IkappaB or  
 CC beta-catenin and is constituted by Skp1 protein, Cull1 protein and a  
 CC complex (SCF complex) of F-box protein containing F-box motif and WD40  
 CC repeat motif and has the amino acid sequence of 45 residues (AAB12813)  
 CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin  
 CC ligase FWD1 protein) and (AAB12813, which is human beta-transducin  
 CC repeat containing protein (beta-Trcp)). The F-box protein can be used for  
 CC the gene therapy of colon cancer by being recombined to a virus vector.  
 XX

SO Sequence 569 AA;  
 XX

Query Match 100.0%; Score 3034; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-286;  
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAEAVIOEKALKFMNSEREDCNNGEPKRIIPKNSLRQTYNSCARLINOETVCLA 60  
 DB 1 MDPAEAVIOEKALKFMNSEREDCNNGEPKRIIPKNSLRQTYNSCARLINOETVCLA 60  
 QY 61 STAMKTENCVAATKILANGSSMIVPKORRLASYEKEKELCYKYPEQMSDOVEFEVHL 120  
 DB 61 STAMKTENCVAATKILANGSSMIVPKORRLASYEKEKELCYKYPEQMSDOVEFEVHL 120  
 QY 121 ISOMCHYOCHGINSYLKPMLODFITLALPARGLDIAENILSYLDKAKSLCAAEIVCKEMY 180  
 DB 121 ISOMCHYOCHGINSYLKPMLODFITLALPARGLDIAENILSYLDKAKSLCAAEIVCKEMY 180  
 QY 181 RYTSOGMLMKKILIERVVRDLSMRGLAERGMQGYLFKNKPPDGNAPNSFYRALYPKII 240  
 DB 181 RYTSOGMLMKKILIERVVRDLSMRGLAERGMQGYLFKNKPPDGNAPNSFYRALYPKII 240  
 QY 241 ODLETIESNMRCGRHSIORHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300  
 DB 241 ODLETIESNMRCGRHSIORHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300  
 QY 301 RILTGHGTVLCLOYDERVITITGSSDSIVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360  
 DB 301 RILTGHGTVLCLOYDERVITITGSSDSIVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360  
 QY 361 VTCSKDRSLAVWDMASPTDITLRRVLVGHRAAVNVDFDDKIVSASGDRITKWNNTSTC 420  
 DB 361 VTCSKDRSLAVWDMASPTDITLRRVLVGHRAAVNVDFDDKIVSASGDRITKWNNTSTC 420  
 QY 421 EFVRTLNGHRRGIACTOYDRRLVVGSSSDNTIRLMDIEGACLRVLEHGEHLVRCIRPDN 480  
 DB 421 EFVRTLNGHRRGIACTOYDRRLVVGSSSDNTIRLMDIEGACLRVLEHGEHLVRCIRPDN 480  
 QY 481 KRIVSGAYDGKIKVMDLVVALDPRAPAGTLCRTLVESHGSRVRLQDFEFOIVSSSHDT 540  
 DB 481 KRIVSGAYDGKIKVMDLVVALDPRAPAGTLCRTLVESHGSRVRLQDFEFOIVSSSHDT 540  
 QY 541 ILTMDFLNDPAQAEPSPRSRTTYTISR 569  
 DB 541 ILTMDFLNDPAQAEPSPRSRTTYTISR 569

RESULT 3  
 ID AAY96697 standard; Protein; 569 AA.  
 AC AAY96697;  
 XX 26-SEP-2000 (first entry)  
 DE Human beta-TrCP.  
 XX E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;  
 KM nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;  
 KM anti-inflammatory; immunosuppressive; cytoskeletal.  
 OS Homo sapiens.  
 XX MO200034447-A2.  
 XX 15-JUN-2000.  
 XX 10-DEC-1999; 99MO-US29371.  
 XX 10-DEC-1998; 98US-0210060.  
 PA (SIGN-) SIGNAL PHARM INC.  
 PA (YISS) YISSUM RES & DEV CO.  
 PI Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A,  
 PI Lavon I, Yaron A;  
 XX WPI: 2000-431294/37.  
 DR N-PSDB; AAS1229.

XX Polyepitope enhancing phosphorylated I-kappa-B ubiquitination useful for  
 PT treating disorder associated with NF-kappa-B activation e.g. cancer,  
 PT comprising amino acid sequence of human E3 ubiquitin ligase or its  
 PT variant  
 PS Claim 21; Page 72-74; 77pp; English.  
 XX Human beta-TrCP, an F-box/WD protein family member, has been shown to  
 CC have homology to human E3 ubiquitin ligase (E3). E3 enhances  
 CC ubiquitination of phosphorylated I-kappa-B, an inhibitor protein of  
 CC nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B  
 CC degradation via the ubiquitin pathway is useful for identifying  
 CC modulators of this process for use in treating diseases associated with  
 CC activation of NF-kappa-B. In vitro analysis suggests that deletion of  
 CC the F-box results in a protein that functions as a dominant negative  
 CC molecule in vivo. Transient over-expression of delta-beta-TrCP (a  
 CC deletion mutant) inhibited the degradation of endogenous I-kappa-B-alpha  
 CC in stimulated Jurkat cells, resulting in accumulation of phosphorylated  
 CC I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B  
 CC activity. E3 and beta-TrCP can be used to modulate NF-kappa-B to treat  
 CC inflammatory diseases, autoimmune diseases, cancer and viral infections.  
 XX Sequence 569 AA:  
 SQ  
 Query Match 100.0%; Score 3034; DB 21; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 1,4e-286;  
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MDPAEAVIOEKALKFMNSEREDCNNGEPKRIIPKNSLRQTYNSCARLINOETVCLA 60  
 DB 1 MDPAEAVIOEKALKFMNSEREDCNNGEPKRIIPKNSLRQTYNSCARLINOETVCLA 60  
 QY 61 STAMKTENCVAATKILANGSSMIVPKORRLASYEKEKELCYKYPEQMSDOVEFEVHL 120  
 DB 61 STAMKTENCVAATKILANGSSMIVPKORRLASYEKEKELCYKYPEQMSDOVEFEVHL 120  
 QY 121 ISOMCHYOCHGINSYLKPMLODFITLALPARGLDIAENILSYLDKAKSLCAAEIVCKEMY 180  
 DB 121 ISOMCHYOCHGINSYLKPMLODFITLALPARGLDIAENILSYLDKAKSLCAAEIVCKEMY 180  
 QY 181 RYTSOGMLMKKILIERVVRDLSMRGLAERGMQGYLFKNKPPDGNAPNSFYRALYPKII 240  
 DB 181 RYTSOGMLMKKILIERVVRDLSMRGLAERGMQGYLFKNKPPDGNAPNSFYRALYPKII 240  
 QY 241 ODLETIESNMRCGRHSIORHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300  
 DB 241 ODLETIESNMRCGRHSIORHCRSETSKGYVCLQYDDQKIVSGLRDNTIKIMDKNTLECK 300  
 QY 301 RILTGHGTVLCLOYDERVITITGSSDSIVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360  
 DB 301 RILTGHGTVLCLOYDERVITITGSSDSIVRVMDVNTGEMLNTLIHCEAVLHLRFNNGMM 360  
 QY 361 VTCSKDRSLAVWDMASPTDITLRRVLVGHRAAVNVDFDDKIVSASGDRITKWNNTSTC 420  
 DB 361 VTCSKDRSLAVWDMASPTDITLRRVLVGHRAAVNVDFDDKIVSASGDRITKWNNTSTC 420  
 QY 421 EFVRTLNGHRRGIACTOYDRRLVVGSSSDNTIRLMDIEGACLRVLEHGEHLVRCIRPDN 480  
 DB 421 EFVRTLNGHRRGIACTOYDRRLVVGSSSDNTIRLMDIEGACLRVLEHGEHLVRCIRPDN 480  
 QY 481 KRIVSGAYDGKIKVMDLVVALDPRAPAGTLCRTLVESHGSRVRLQDFEFOIVSSSHDT 540  
 DB 481 KRIVSGAYDGKIKVMDLVVALDPRAPAGTLCRTLVESHGSRVRLQDFEFOIVSSSHDT 540  
 QY 541 ILTMDFLNDPAQAEPSPRSRTTYTISR 569  
 DB 541 ILTMDFLNDPAQAEPSPRSRTTYTISR 569

RESULT 4  
 ID AAY83041 standard; Protein; 569 AA.  
 AC AAY83041;  
 XX 26-SEP-2000 (first entry)  
 DE Human beta-TrCP.  
 XX E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;  
 KM nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;  
 KM anti-inflammatory; immunosuppressive; cytoskeletal.  
 OS Homo sapiens.  
 XX MO200034447-A2.  
 XX 15-JUN-2000.  
 XX 10-DEC-1999; 99MO-US29371.  
 XX 10-DEC-1998; 98US-0210060.  
 PA (SIGN-) SIGNAL PHARM INC.  
 PA (YISS) YISSUM RES & DEV CO.  
 PI Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A,  
 PI Lavon I, Yaron A;  
 XX WPI: 2000-431294/37.  
 DR N-PSDB; AAS1229.

XX AAY63041;  
 AC  
 DT 16-AUG-2000 (first entry)  
 XX  
 XX F-box protein FBP-1.  
 XX  
 XX F-box protein; FBP; diagnosis; treatment; screening; agonist;  
 KW antagonist; proliferative disorder; differentiative disorder;  
 KW breast cancer; prostate cancer; ovarian cancer; cancer;  
 KW small cell lung carcinoma; immune disorder; cardiovascular disorder;  
 KW inflammatory disorder; human.  
 OS  
 OS Homo sapiens.  
 XX  
 XX WO200012679-A1.  
 PN  
 XX  
 XX 09-MAR-2000.  
 PD  
 XX  
 XX 27-AUG-1999; 99WO-US19560.  
 PF  
 XX  
 XX 28-AUG-1998; 98US-0098335.  
 PR 03-FEB-1999; 99US-0118558.  
 PR 15-MAR-1999; 99US-0124449.  
 XX  
 XX (UYNY ) UNIV NEW YORK STATE.  
 PA  
 XX  
 XX Chlaour DS, Pagano M, Latres E;  
 PI  
 XX WPI: 2000-256635/22.  
 DR N-PSDB; MAZ93350.  
 DR  
 XX Novel nucleic acid for screening compounds useful for treating  
 PT proliferative and differentiative disorders such as cancer and immune  
 PT disorders comprises sequences encoding ubiquitin ligases -  
 PS  
 PS Disclosure; Figure 3a; 245pp; English.  
 XX  
 XX Nucleic acids encoding substrate-targeting subunits of ubiquitin  
 CC ligases with F-box motifs (F-box proteins) are useful for diagnosis  
 CC of proliferative and differentiated related disorders by measuring  
 CC FBP gene expression. Cells expressing such proteins or  
 CC their fragments are useful for screening compounds. The compounds  
 CC are agonists or antagonists, which are useful for treating a  
 CC proliferative or differentiative disorder in a mammal such as  
 CC breast, ovarian and prostate cancer and small cell lung carcinoma  
 CC and also major opportunistic infections, immune disorders,  
 CC cardiovascular diseases and inflammatory disorders. FBP protein,  
 CC analogs, derivatives and their subsequences, anti-FBP antibodies  
 CC are also useful in diagnosis of the disorders.  
 XX  
 XX Sequence 569 AA;  
 XX

QY	241	ODIEIEENMGRGRIStORICRSTSGVYCLQYDDOKIYSGRLDNTIKIMKNTLECK	300
Db	241	QDIETIESMWRGRIStORICRSTSGVYCLQYDDOKIYSGRLDNTIKIMKNTLECK	3000
QY	301	RLTNGHTSSVLCLOQYDEVIIITGSSDSIVRWADVNTGEMLNTLIHCEAVLHLRFNNGMM	3600
Db	301	RLTNGHTSSVLCLOQYDEVIIITGSSDSIVRWADVNTGEMLNTLIHCEAVLHLRFNNGMM	3600
QY	361	WYCSKDRSIAWMDMSPDITLRLVLYGHRAAVNVDPDXYIYASAGDPTIKWMMSTC	4200
Db	361	WYCSKDRSIAWMDMSPDITLRLVLYGHRAAVNVDPDXYIYASAGDPTIKWMMSTC	4200
QY	421	EFVRLNHRKRGICLOQYRDRLVYSSGSDNTIRLMDIECGACLRVLEGHELYRCIRFDN	4800
Db	421	EFVRLNHRKRGICLOQYRDRLVYSSGSDNTIRLMDIECGACLRVLEGHELYRCIRFDN	4800
QY	481	KRIYSGAYDGKIKWMDLVAAALDPRAPAGTLCRLTVBHSGRVRFLODFEQIYSSHDDY	5400
Db	481	KRIYSGAYDGKIKWMDLVAAALDPRAPAGTLCRLTVBHSGRVRFLODFEQIYSSHDDY	5400
QY	541	ILIMDELNDPAAQAEPSPSPRTTYISR	5600
Db	541	ILIMDELNDPAAQAEPSPSPRTTYISR	5600

RESULT 5	
AAV83250	
ID	AAV83250 standard; Protein; 569 AA
XY	

DT 16-AUG-2000 (first entry)  
 VV

F-box protein hbetatrcp

**KW** ubiquitin ligase; SCF; F-box protein; targeted degradation;

KW oncprotein; Huntington's disease; gene knockout; delivery systems;

XX  
MV  
JUN 1968  
XX

XX Homo sapiens.

PN WO2000022110-  
XX

PD 20-APR-2000.  
yy

PF 08-OCT-1999; 99WO-US23705.

PR 09-OCT-1998; 98US-0103787.

PA (HARD ) HARVARD COLLEGE

AA      Zhou P, Howley P:  
PI

XX ,  
DR WPT. 2000-317970/27

DR N-PSDB; AAZ93710.  
XY

PT Targeting degradation

ubiquitin protein 15

XX

STAT 3, Page 1/1, 1  
XX

CC The F-box proteins  
CC liases) which can

polypeptide in vivo.  
the ubiquitin-ligase

CC the target polypeptide

CC increasing the level

CC mediated proteolysis

PT Targeting degradation of polypeptide useful for treating cancer and  
PT other proliferative disorders, involves conjugating polypeptide with  
PT ubiquitin protein ligase or inhibiting ubiquitination using organic  
PT compound  
XX  
XX  
PS Claim 9, page 171; 185pp; English.  
PS  
XX  
XX The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin  
CC ligases) which can be used for the targeted degradation of a target  
CC polypeptide in vivo. Targeted degradation is achieved by expressing  
CC the ubiquitin ligase in a cell linked to the interaction domain of  
CC the target polypeptide and thereby recruiting the target polypeptide  
CC to the ubiquitin ligase. Such methods are useful for decreasing or  
CC increasing the level of a target polypeptide and for creating and  
CC expressing a destabilized polypeptide which is subjected to SCF  
CC mediated proteolysis. Degrading any desired protein in a cell is

useful for preventing or treating diseases caused by the presence of abnormal amount of the specific polypeptides, for drug discovery and for gene therapy. Diseases treated include cancer, by degradation of oncoproteins, Huntington's disease, other proliferative disorders and microbial infections. The method provides a quick and easy alternative to gene knockout technology. The target polypeptide can be degraded at all stages, or a specific stage, of development in the mature animal.

**SQ Sequence 569 AA;**

Query Match	100.0%;	Score 3034;	DB 21;	length 569;
Best Local Similarity	100.0%;	Pred. No. 1.4e-286;		
Matches 569;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	MDPEAVYLOEALKEFNMSSEEDCCNNNEPPKIIIPKENSILROTYNSCARCLJNOEYVCLA	60
Dd	1	MDPEAVYLOEALKEFNMSSEEDCCNNNEPPKIIIPKENSILROTYNSCARCLJNOEYVCLA	60
Qy	61	STAKTENCYAKTKLIANGTSSMIVPKORKLSASYEKKEKELCVKYFEQWSSDVEFEVHL	120
Dd	61	STAKTENCYAKTKLIANGTSSMIVPKORKLSASYEKKEKELCVKYFEQWSSDVEFEVHL	120
Qy	121	ISOMCHTOHGHIINSTYKPMLODFITPLPARGLDHIABNLLSYLDKASLCAELVCKEYT	180
Dd	121	ISOMCHTOHGHIINSTYKPMLODFITPLPARGLDHIABNLLSYLDKASLCAELVCKEYT	180
Qy	181	RYTSGMIMKTLIRMYRTDSLIMGLEBRGMOGYLEKKNKPPGCMAPNPSFYALYPKII	240
Dd	181	RYTSGMIMKTLIRMYRTDSLIMGLEBRGMOGYLEKKNKPPGCMAPNPSFYALYPKII	240
Qy	241	QDIEETIESNMNRCKGHRSLORIHCRSEFSKGYVCLQYDDQKTVSGLRDNTIKIWKDKNTLECK	300
Dd	241	QDIEETIESNMNRCKGHRSLORIHCRSEFSKGYVCLQYDDQKTVSGLRDNTIKIWKDKNTLECK	300
Qy	301	RLNGHTGSLVCLQYDERVLIITGSSDSTVRWYDVNTGEMLNTLIIHCEAVLHLRFNNGM	360
Dd	301	RLNGHTGSLVCLQYDERVLIITGSSDSTVRWYDVNTGEMLNTLIIHCEAVLHLRFNNGM	360
Qy	361	WYCSMDRSIAWMDASPFDDITLRRVLVGHRAAVWVDPDDKYIYSASGDPRTIKYWNSTC	420
Dd	361	WYCSMDRSIAWMDASPFDDITLRRVLVGHRAAVWVDPDDKYIYSASGDPRTIKYWNSTC	420
Qy	421	EFVFTLNGHKRGICLQYRDRLVYSSGSSDNITRLMDIECGACLRVLBGHELVYCIFFDN	480
Dd	421	EFVFTLNGHKRGICLQYRDRLVYSSGSSDNITRLMDIECGACLRVLBGHELVYCIFFDN	480
Qy	481	KRIYSGAVDGKIKYWDILVAALDPRAPACTLCLRTLVEHSGRVFLQDFEFOIYSSSHDT	540
Dd	481	KRIYSGAVDGKIKYWDILVAALDPRAPACTLCLRTLVEHSGRVFLQDFEFOIYSSSHDT	540
Qy	541	ILIMDFLNDPAAQAEPPSPSRITTYISR	569
Dd	541	ILIMDFLNDPAAQAEPPSPSRITTYISR	569

RESULT 6  
AAV44249  
ID AAV44249 standard; Protein; 569 AA.

AC AAY44249

DT 28-FEB-2000 (first entry)

DE Human cell signalling protein-12.

KW Cell signalling protein-12; CSRP-12; cell proliferation; inflammation; disorder; diabetes; cancer; hepatitis; AIDS

KW arteriosclerosis; Addison's disease; multiple sclerosis.

OS Homo sapiens.

**FH Key**

FT	Modified-site	19	"Potential phosphorylation site"
FT	Modified-site	39	"Potential phosphorylation site"
FT	Modified-site	91	"Potential phosphorylation site"
FT	Modified-site	109	"Potential phosphorylation site"
FT	Modified-site	162	"Potential phosphorylation site"
FT	Modified-site	266	"Potential phosphorylation site"
FT	Modified-site	288	"Potential phosphorylation site"
FT	Modified-site	328	"Potential phosphorylation site"
FT	Modified-site	376	"Potential phosphorylation site"
FT	Modified-site	381	"Potential phosphorylation site"
FT	Modified-site	411	"Potential phosphorylation site"
FT	Modified-site	418	"Potential phosphorylation site"
FT	Modified-site	451	"Potential phosphorylation site"
FT	Modified-site	514	"Potential phosphorylation site"
FT	Modified-site	519	"Potential phosphorylation site"
FT	Modified-site	535	"Potential phosphorylation site"
FT	Modified-site	536	"Potential phosphorylation site"
FT	Modified-site	17	"Potential phosphorylation site"
FT	Modified-site	77	"Potential glycosylation site"
FT	Modified-site	416	"Potential glycosylation site"
FT	Modified-site	320..334	"Potential glycosylation site"
FT	Region	360..374	/label= Signature_sequence
FT	Region	403..417	/label= Signature_sequence
FT	Region	443..457	/label= Signature_sequence
FT	Region	483..497	/label= Signature_sequence
FT	Region	532..546	/label= Signature_sequence
FT	Region		/label= Signature_sequence
XX			
PN	WO958558-A2.		
XX			
PD	18-NOV-1999.		
XX			
PF	13-MAY-1999;	99WO-US10567.	
XX			
PR	13-MAY-1998;	98US-0085343.	
XX	26-AUG-1998;	98US-0098010.	
XX			
PA	(INCY- ) INCYTE PHARM INC.		
XX			
PI	Bandman O, Hillman JL, Lai P, Yue H, Tang YT, Patterson C,		
XX	Baughn MR, Yang J;		
DR	WPI, 2000-086432/07.		
XX	N-PSDB; AA229233.		
PT	Human cell signaling proteins useful for, e.g. diagnosing cell		
XX	proliferative and inflammatory disorders		
XX			

PS Claim 1; Page 77-78; 90pp; English.

CC The present sequence is cell signalling protein-12 (CSIGP-12) encoded  
CC by cDNA obtained from Inocyte clone 3239149 of COLADU001 library. It is  
CC expressed in musculoskeletal, gastrointestinal and nervous tissues and is  
CC found to be homologous to beta-transducin repeats containing  
CC protein. Fragments of CSIGP encoding nucleic acid can be used as  
CC hybridisation probe for detecting CSIGP related sequences or allelic  
CC variants. Recombinant CSIGP can be produced in host cells by transforming  
CC them with genetically engineered vectors. Agonists or antagonists can be  
CC used in the treatment of cell proliferative and inflammatory disorders  
CC associated with decreased or increased CSIGP expression. CSIGP is used in  
CC the diagnosis, prevention and treatment of cell proliferative disorders  
CC like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory  
CC disorders like AIDS, Addison's disease, multiple sclerosis, etc.

CC Sequence 569 AA:

Query Match 100.0%; Score 3034; DB 21; Length 569;  
Best Local Similarity 100.0%; Pred. No. 1,4e-286;  
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAAVAVLQEKALFKFNSREDONNGEPKRIIPKNSLRQTYNSCARCLNQEYVCLA 60  
DB 1 MDPAAVAVLQEKALFKFNSREDONNGEPKRIIPKNSLRQTYNSCARCLNQEYVCLA 60  
QY 61 STAMKTEVCVAKTKLANGTSSMIVPKOKRLSASYEKEKELCVKFFQMSDQVEFEHL 120  
DB 61 STAMKTEVCVAKTKLANGTSSMIVPKOKRLSASYEKEKELCVKFFQMSDQVEFEHL 120  
QY 121 ISQCHYGHGHIINSYKLPMLQDRFETALPARGLDHAENILSYDAKSLCAAEIVCKEMY 180  
DB 121 ISQCHYGHGHIINSYKLPMLQDRFETALPARGLDHAENILSYDAKSLCAAEIVCKEMY 180  
QY 181 RVTSDGMLMKLLIERMVRTDSLMRGLAERNGQGYLFKNKPPDGNAPNSFYRALPKIT 240  
DB 181 RVTSDGMLMKLLIERMVRTDSLMRGLAERNGQGYLFKNKPPDGNAPNSFYRALPKIT 240  
QY 241 ODIEETIESNMRCGRHSIORHCRSETSKGYVCIQYDDOKIVSGLRNTIKIMDKNTLECK 300  
DB 241 ODIEETIESNMRCGRHSIORHCRSETSKGYVCIQYDDOKIVSGLRNTIKIMDKNTLECK 300  
QY 301 RIITGHGTVLCIQYDERVITITGSSDSTVRWVDVNTGEMLNTLIHCEAVLHLRFNNGM 360  
DB 301 RIITGHGTVLCIQYDERVITITGSSDSTVRWVDVNTGEMLNTLIHCEAVLHLRFNNGM 360  
QY 361 VTCSKDRSIAVWMAASPTDITLRRLVVGHRAAVNVDFDDKIYVSASGDRITKVMNTSTC 420  
DB 361 VTCSKDRSIAVWMAASPTDITLRRLVVGHRAAVNVDFDDKIYVSASGDRITKVMNTSTC 420  
QY 421 EFVRTLNGHKGRTACIQYDRDLVVGSSDNTIRLMDIEGACLRVLEGHEELVRCIRFPN 480  
DB 421 EFVRTLNGHKGRTACIQYDRDLVVGSSDNTIRLMDIEGACLRVLEGHEELVRCIRFPN 480  
QY 481 KRIVSAGYOGKITVMDLVVALDPRAGTLCRTLVESHGGRVFRIDFQIYVSSSHDT 540  
DB 481 KRIVSAGYOGKITVMDLVVALDPRAGTLCRTLVESHGGRVFRIDFQIYVSSSHDT 540  
QY 541 ILITMDFLNDPAQAQEPSPSRITYTISR 569  
DB 541 ILITMDFLNDPAQAQEPSPSRITYTISR 569

RESULT 7

AAAB48298  
ID AAAB48298 standard; protein; 569 AA.

AC AAAB48298;

DT 02-APR-2001 (first entry)

DE Human Zf11 protein.

XX

KW S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;  
KW CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;  
KW Bad; Bcl-2; tumour; cytosolic.

OS Homo sapiens.

PN WO200075184-A1.

PD 14-DEC-2000.

PF 05-JUN-2000; 2000WO-US15449.

PR 04-JUN-1999; 99US-0137494.

PA (UYUA ) UNIV YALE.

PI Zhang H, Tsvetkov LM, Kondo T;

DR WPI: 2001-061703/07.

DR N-PSDB: AAC84610.

PT Modulating polypeptide levels in a cell, diagnosing and treating tumor,  
PT involves altering levels of proteins such as S-phase kinase associated  
PT proteins 1, 2 and cullin/CDC53 proteins -

PS Claim 3; Page 130-132; 162pp; English.

CC The invention relates to methods of altering the polypeptide levels in a  
CC cell, using proteins selected from S-phase kinase associated proteins 1  
CC and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the  
CC cullin/CDC53 family of proteins). The method is useful for altering the  
CC level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2  
CC polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for  
CC detecting tumours, and in monitoring tumor treatment in a mammal. Agents  
CC that modulate interactions between SKP and target proteins are useful for  
CC treating tumours.

CC Sequence 569 AA:

Query Match 99.8%; Score 3027; DB 22; Length 569;  
Best Local Similarity 99.8%; Pred. No. 6.9e-286;  
Matches 568; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDPAAVAVLQEKALFKFNSREDONNGEPKRIIPKNSLRQTYNSCARCLNQEYVCLA 60  
DB 1 MDPAAVAVLQEKALFKFNSREDONNGEPKRIIPKNSLRQTYNSCARCLNQEYVCLA 60  
QY 61 STAMKTEVCVAKTKLANGTSSMIVPKOKRLSASYEKEKELCVKFFQMSDQVEFEHL 120  
DB 61 STAMKTEVCVAKTKLANGTSSMIVPKOKRLSASYEKEKELCVKFFQMSDQVEFEHL 120  
QY 121 ISQCHYGHGHIINSYKLPMLQDRFETALPARGLDHAENILSYDAKSLCAAEIVCKEMY 180  
DB 121 ISQCHYGHGHIINSYKLPMLQDRFETALPARGLDHAENILSYDAKSLCAAEIVCKEMY 180  
QY 181 RVTSDGMLMKLLIERMVRTDSLMRGLAERNGQGYLFKNKPPDGNAPNSFYRALPKIT 240  
DB 181 RVTSDGMLMKLLIERMVRTDSLMRGLAERNGQGYLFKNKPPDGNAPNSFYRALPKIT 240  
QY 241 ODIEETIESNMRCGRHSIORHCRSETSKGYVCIQYDDOKIVSGLRNTIKIMDKNTLECK 300  
DB 241 ODIEETIESNMRCGRHSIORHCRSETSKGYVCIQYDDOKIVSGLRNTIKIMDKNTLECK 300  
QY 301 RIITGHGTVLCIQYDERVITITGSSDSTVRWVDVNTGEMLNTLIHCEAVLHLRFNNGM 360  
DB 301 RIITGHGTVLCIQYDERVITITGSSDSTVRWVDVNTGEMLNTLIHCEAVLHLRFNNGM 360  
QY 361 VTCSKDRSIAVWMAASPTDITLRRLVVGHRAAVNVDFDDKIYVSASGDRITKVMNTSTC 420  
DB 361 VTCSKDRSIAVWMAASPTDITLRRLVVGHRAAVNVDFDDKIYVSASGDRITKVMNTSTC 420  
QY 421 EFVRTLNGHKGRTACIQYDRDLVVGSSDNTIRLMDIEGACLRVLEGHEELVRCIRFPN 480  
DB 421 EFVRTLNGHKGRTACIQYDRDLVVGSSDNTIRLMDIEGACLRVLEGHEELVRCIRFPN 480

DB 421 EFVRLNGHRRGIAQLQYRDRLVVGSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDN 480  
 QY 481 KRIVSGAYDGRKIKVMDLVALDPRAPAGTLCRLTVEHSGRVRLQDFEQIVSSSHDPT 540  
 DB 481 KRIVSGAYDGRKIKVMDLVALDPRAPAGTLCRLTVEHSGRVRLQDFEQIVSSSHDPT 540  
 QY 541 ILIMDFLNDPAAQAEPPRSPRTTYTISR 569  
 DB 541 ILIMDFLNDPAAQAEPPRSPRTTYTISR 569  
 RESULT 8  
 AAM00960  
 ID AAM00960 standard; Protein; 608 AA.  
 AC AAM00960;  
 XX 01-OCT-2001 (first entry)  
 DT  
 XX Human bone marrow protein, SEQ ID NO: 436.  
 DE  
 XX Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;  
 KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;  
 KW immunosuppressive; gene therapy; cytokine cell proliferation;  
 KW cell differentiation modulator; immune disorder; infection; cancer;  
 KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.  
 XX Homo sapiens.  
 OS  
 XX  
 PN W0200153453-A2.  
 PD 26-JUL-2001.  
 XX  
 PE 23-DEC-2000; 2000WO-US34960.  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 30-NOV-2000; 2000US-0250583.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 PI Ford JB, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;  
 PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Drmanac RT;  
 DR WPI; 2001-488707/53.  
 DR N-PSDB; AAH90079.  
 PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful  
 PT for treating e.g. cancer and immune deficiency disorders -  
 PS Claim 10; Page 523-524; 648pp; English.  
 CC The present sequence is one of 251 novel human polypeptides encoded  
 CC by a bone marrow-expressed polynucleotide. The polynucleotide and the  
 CC polypeptide encoded by it are useful in the treatment of various  
 CC immune deficiencies and disorders. The deficiencies and disorders may  
 CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal  
 CC infection, or may result from an autoimmune disorder, a coagulation  
 CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,  
 CC suppression of an inflammatory response or treatment of a nervous  
 CC system disorder such as Alzheimer's disease. Detection of the presence  
 CC or increased expression of the polynucleotide or the protein it  
 CC encodes is useful for the diagnosis and/or prognosis of one  
 CC or more types of cancer. The polynucleotide and polypeptide can be  
 CC used as nutritional sources or supplements and in the screening of  
 CC chemical compounds as potential drugs.

SQ Sequence 608 AA:  
 Query Match 99.8%; Score 3027; DB 22; Length 608;  
 Best Local Similarity 99.6%; Pred. No. 7.7e-286;  
 Matches 567; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MDPAEAVLQEKALKFENSSEREDCNGEPPRKIIPEKNSLRQTVSCARLCLNDQTVCLA 60  
 DB 40 MDPEAVLQEKALKFENSSEREDCNGEPPRKIIPEKNSLRQTVSCARLCLNDQTVCLA 99  
 QY 61 STAMKTEHCVAATKTLANGSSMTIVPRORLASAYEKKEKLVKYPEFOWSESQVFEVHL 120  
 DB 100 STAMKTEHCVAATKTLANGSSMTIVPRORLASAYEKKEKLVKYPEFOWSESQVFEVHL 159  
 QY 121 ISOMCHYOHGHINSYLPMLQDRFTALPARGLDHAENILSYLDKSLCAAEVCKEMV 180  
 DB 160 ISOMCHYOHGHINSYLPMLQDRFTALPARGLDHAENILSYLDKSLCAAEVCKEMV 219  
 QY 181 RYTSDEMLMKKLIERNVPTDLSIRGLAERRGQYIFKPKPPDGNAPPSFTBALYPKII 240  
 DB 220 RYTSDEMLMKKLIERNVPTDLSIRGLAERRGQYIFKPKPPDGNAPPSFTBALYPKII 279  
 QY 241 QDIETTESNMRCGRHSLOTHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECK 300  
 DB 280 QDIETTESNMRCGRHSLOTHCRSETSKGVYCLQYDDOKIVSGLRDNTIKIMDKNTLECK 339  
 QY 301 RILTGHGTVLCLOYDERVYITGSSDSSTVAVDVTGEMLNTLIHRCFAVLHLPFNCGM 360  
 DB 340 RILTGHGTVLCLOYDERVYITGSSDSSTVAVDVTGEMLNTLIHRCFAVLHLPFNCGM 399  
 QY 361 VTCSKDRSLAVWDMASPTDITLRLVYVGHRAAVNYDFFDDKITYVSGSRTIKVWNTSTC 420  
 DB 400 VTCSKDRSLAVWDMASPTDITLRLVYVGHRAAVNYDFFDDKITYVSGSRTIKVWNTSTC 459  
 QY 421 EFVRLNGHRRGIAQLQYRDRLVVGSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDN 480  
 DB 460 EFVRLNGHRRGIAQLQYRDRLVVGSSDNTIRLMDIEGACLRVLEGHEELVRCIRFDN 519  
 QY 481 KRIVSGAYDGRKIKVMDLVALDPRAPAGTLCRLTVEHSGRVRLQDFEQIVSSSHDPT 540  
 DB 520 KRIVSGAYDGRKIKVMDLVALDPRAPAGTLCRLTVEHSGRVRLQDFEQIVSSSHDPT 579  
 QY 541 ILIMDFLNDPAAQAEPPRSPRTTYTISR 569  
 DB 580 ILIMDFLNDPAAQAEPPRSPRTTYTISR 608  
 RESULT 9  
 AAM78582  
 ID AAM78582 standard; Protein; 605 AA.  
 AC AAM78582;  
 XX 06-NOV-2001 (first entry)  
 DT  
 XX Human protein SEQ ID NO 1244.  
 DE  
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation.  
 XX Homo sapiens.  
 OS  
 XX  
 PN W0200157190-A2.  
 PD 09-AUG-2001.  
 XX  
 PD 05-FEB-2001; 2001WO-US04098.  
 PF 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.  
 PR 01-SEP-2000; 2000US-0654936.  
 PR 15-SEP-2000; 2000US-0663561.  
 PR 20-OCT-2000; 2000US-0693325.  
 PR 30-NOV-2000; 2000US-0728422.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wehrman T, Goodrich R;  
 XX  
 DR WPI: 2001-476283/51.  
 DR N-PSDB; AAK51715.  
 PT Nucleic acids encoding polypeptides with cytokine-like activities,  
 PT useful in diagnosis and gene therapy -  
 XX  
 PS Claim 20; Page 3503-3504; 6221pp; English.  
 XX  
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activity/inhibn activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAM80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.  
 CC  
 XX Sequence 605 AA:  
 SO  
 Query Match 99.1%; Score 3006; DB 22; Length 605;  
 Best Local Similarity 94.0%; Pred. No. 8.5e-284;  
 Matches 569; Conservative 0; Mismatches 0; Indels 36; Gaps 1;  
 OY 1 MDPAEAVLOEKALK-----FNNSSREDC 24  
 DB 1 MDPAEAVLOEKALKFKMCSMPRLIWGSSSLADSMPSLRCLNPGTALTAFAFNSSREDC 60  
 OY 25 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAKTENCYAKTKLANGTSSMIV 84  
 DB 61 NNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCLASTAKTENCYAKTKLANGTSSMIV 120  
 OY 85 PROKRISASYEKEKELCVYFFOWSESDOVEVEHLISOMCHYOHGHSINSLKPMLOQDF 144  
 DB 121 PROKRISASYEKEKELCVYFFOWSESDOVEVEHLISOMCHYOHGHSINSLKPMLOQDF 180  
 OY 145 ITALPARGDHIAENILSVLDAKSICAAELVCKEYRYATSDMKLIERVNRDLSMR 204  
 DB 181 ITALPARGDHIAENILSVLDAKSICAAELVCKEYRYATSDMKLIERVNRDLSMR 240  
 OY 205 GLAERRGMOQYLFKNKPPDGNAPNSFYRALPKIIOETIESNMGCRHSIQRIHCRS 264  
 DB 241 GLAERRGMOQYLFKNKPPDGNAPNSFYRALPKIIOETIESNMGCRHSIQRIHCRS 300  
 OY 265 ETSKGVYCLQYDDOKIVSLGRNTIKIMDKNTLECKRITLTGHTGVLQYDERVITIGS 324  
 DB 301 ETSKGVYCLQYDDOKIVSLGRNTIKIMDKNTLECKRITLTGHTGVLQYDERVITIGS 360  
 OY 325 SDSIVRVMDVTGEMILNLIHCEAVLHLRFNNGMAYTCSKDRSTAVYDMSPTDITLR 384  
 DB 361 SDSIVRVMDVTGEMILNLIHCEAVLHLRFNNGMAYTCSKDRSTAVYDMSPTDITLR 420  
 OY 385 VLVGRRAAVNVVDPDKTIYVSAGDRTIKVNTSTCEFEVRLNKGKRIACIQYRDRLVV 444  
 DB 421 VLVGRRAAVNVVDPDKTIYVSAGDRTIKVNTSTCEFEVRLNKGKRIACIQYRDRLVV 480

OY 445 SGSSDNTIRLMDIEGACLRVEGHEELVRCIRFNNKRIYSGAYDGKIKWMDLVAALDP 504  
 DB 481 SGSSDNTIRLMDIEGACLRVEGHEELVRCIRFNNKRIYSGAYDGKIKWMDLVAALDP 540  
 OY 505 APAGTLCRLTVLHSGRVRLQFDEFQIVSSSHDPTILMDLNDPAAQAEPPSPSRXY 564  
 DB 541 APAGTLCRLTVLHSGRVRLQFDEFQIVSSSHDPTILMDLNDPAAQAEPPSPSRXY 600  
 OY 565 TYISR 569  
 DB 601 TYISR 605  
 RESULT 10  
 AAB12812  
 ID AAB12812 standard; protein; 569 AA.  
 XX  
 AC AAB12812;  
 XX  
 DT 27-NOV-2000 (first entry)  
 XX  
 DE Mouse ubiquitin ligase FWD1 protein SEQ ID NO.2.  
 XX  
 KW Ubiquitin ligase SCF complex; F-box protein; ubiquitination; Ikappab;  
 KW beta-catenin; Skp1; Cull; F-box motif; WD40 repeat motif; FWD1;  
 KW gene therapy; colon cancer; beta-transducin repeat containing protein;  
 KW beta-TRCP.  
 XX  
 OS Mus musculus.  
 XX  
 PN JP2000166542-A.  
 PD 20-JUN-2000.  
 XX  
 XX 02-DEC-1998; 98JP-0343437.  
 XX  
 PR 02-DEC-1998; 98JP-0343437.  
 XX  
 PA (KAGA-) KAGACU GIUTTSU SHINKO JIGYODAN.  
 XX  
 DR WPI: 2000-485550/43.  
 DR N-PSDB; AAA73131.  
 XX  
 PT F-box protein of ubiquitin ligase SCF complex which promotes the  
 PT ubiquitination of Ikappab or beta-catenin -  
 PS  
 PS Claim 2; Page 9-10; 19pp; Japanese.  
 CC The present invention describes an F-box motif protein of ubiquitin  
 CC ligase SCF complex which promotes the ubiquitination of Ikappab or  
 CC beta-catenin and is constituted by Skp1 protein, Cull protein and a  
 CC complex (SCF complex) of F-box protein containing F-box motif and a  
 CC repeat motif and has the amino acid sequence of 45 residues (AAB12811)  
 CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin  
 CC ligase FWD1 protein) and (AAB12813, which is human beta-transducin  
 CC repeat containing protein (beta-TRCP)). The F-box protein can be used for  
 CC the gene therapy of colon cancer by being recombined to a virus vector.  
 CC  
 XX Sequence 569 AA:  
 SO  
 Query Match 98.8%; Score 2997; DB 21; Length 569;  
 Best Local Similarity 98.6%; Pred. No. 5.8e-283;  
 Matches 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 OY 1 MDPAEAVLOEKALKFNNSSREDCNNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCL 60  
 DB 1 MDPAEAVLOEKALKFNNSSREDCNNGEPPRKTIPEKNSLRQTYNSCARLCLNOETVCL 60  
 OY 61 STAMKTENCYAKTKLANGTSSMTVPKORISASYEKEKELCVYFFOWSESDOVEVEHL 120  
 DB 61 STAMKTENCYAKTKLANGTSSMTVPKORISASYEKEKELCVYFFOWSESDOVEVEHL 120  
 OY 121 ISOMCHYOHGHSINSLKPMLOQDFITALPARGDHIAENILSVLDAKSICAAELVCKEY 180

```

Db 121 ISQCHYOHGHINSYLPMLQDRDITLPAKGLDHIENILSYLDAKSLCAAEIVCKEWM 180
OY 181 RYVSDGMLMKKLLERWRTDLSLWGLAERKMGQYLFKNPPGNAFPNSFYALYPKII 240
Db 181 RYVSDGMLMKKLLERWRTDLSLWGLAERKMGQYLFKNPPGNAFPNSFYALYPKII 240
OY 241 ODITISNMRCGRHSIQRHCHRETSKGYCLOYDDOKIVSGLRDNTIKIMDKSTLECK 300
Db 241 ODITISNMRCGRHSIQRHCHRETSKGYCLOYDDOKIVSGLRDNTIKIMDKSTLECK 300
OY 301 RILGHTGSVLCLOYDERVYITGSSDSTVRVWVNTGEMNTLIHCEAVLHLPFNNGM 360
Db 301 RILGHTGSVLCLOYDERVYITGSSDSTVRVWVNTGEMNTLIHCEAVLHLPFNNGM 360
OY 361 VTCKSDRSIAVMDASPTDITLRVLYGHRAAVNVDFDKYIVSASGDTIKYMNSTC 420
Db 361 VTCKSDRSIAVMDASPTDITLRVLYGHRAAVNVDFDKYIVSASGDTIKYMNSTC 420
OY 421 EFVFTLNGHKGIGIACLOYRDLVYVSSGSDNTIRLMDIEGACLVLEGHELVRCIRFDN 480
Db 421 EFVFTLNGHKGIGIACLOYRDLVYVSSGSDNTIRLMDIEGACLVLEGHELVRCIRFDN 480
OY 481 KRIVSGAYDGKIKVMDLVAAIDPPAPAGTCLRTLVHSGRVFLQDFEPOIVSSSHDT 540
Db 481 KRIVSGAYDGKIKVMDLVAAIDPPAPAGTCLRTLVHSGRVFLQDFEPOIVSSSHDT 540
OY 541 ILIMDFLNDPAAQAEPSPSRITYTISR 569
Db 541 ILIMDFLNDPAAHAEPSPSRITYTISR 569

RESULT 11
AAV83254
ID AAV83254 standard; Protein: 569 AA.
AC
XX AAV83254;
XX
XX 16-AUG-2000 (first entry)
XX
XX F-box protein FMDIP.
XX
XX Ubiquitin ligase; SCF; F-box protein; targeted degradation;
XX destabilization; proteolysis; drug discovery; gene therapy; cancer;
XX oncoprotein; Huntington's disease; gene knockout; delivery systems;
XX mouse; ss.
XX
XX Mus musculus.
XX
XX WO200022110-A2.
XX
XX 20-APR-2000.
XX
XX 08-OCT-1999; 99WO-US23705.
XX
XX 09-OCT-1998; 98US-0103787.
XX
XX (HARD) HARVARD COLLEGE.
XX
XX Zhou P, Howley P;
XX
XX WPI; 2000-317970/27.
XX
XX N-PSDB; AA293714.
XX
XX Targeting degradation of polypeptide useful for treating cancer and
XX other proliferative disorders; involves conjugating polypeptide with
XX ubiquitin protein ligase or inhibiting ubiquitination using organic
XX compound
XX
XX Claim 9; Page 184-185; 185pp; English.
XX
XX The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
XX ligases) which can be used for the targeted degradation of a target

```

```

CC polypeptide in vivo. Targeted degradation is achieved by expressing
CC the ubiquitin ligase in a cell linked to the interaction domain of
CC the target polypeptide and thereby recruiting the target polypeptide
CC to the ubiquitin ligase. Such methods are useful for decreasing or
CC increasing the level of a target polypeptide and for creating and
CC expressing a destabilized polypeptide which is subjected to SCF
CC mediated proteolysis. Degrading any desired protein in a cell is
CC useful for preventing or treating diseases caused by the presence of
CC abnormal amount of the specific polypeptides, for drug discovery and
CC for gene therapy. Diseases treated include cancer, by degradation of
CC oncoproteins, Huntington's disease, other proliferative disorders and
CC microbial infections. The method provides a quick and easy
CC alternative to gene knockout technology. The target polypeptide can
CC be degraded at all stages, or a specific stage, of development in the
CC mature animal.
XX
XX Sequence 569 AA:
SQ
Query Match 98.8%; Score 2997; DB 21; Length 569;
Best local Similarity 98.6%; Pred. No. 5,8e-283;
Matches 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
OY 1 MDPAAEVLQEKALKFNNSSREDCCNNGEPPRKIIPEKNSLRQYNSCARLCTNOETVCLA 60
Db 1 MDPAAEVLQEKALKFNNSSREDCCNNGEPPRKIIPEKNSLRQYNSCARLCTNOETVCLA 60
OY 61 STAMKTEVCYAKTKLANGTSSMIVPQKRLSASEKEKELCYVFEQWSESDOVEVEHL 120
Db 61 STAMKTEVCYAKTKLANGTSSMIVPQKRLSASEKEKELCYVFEQWSESDOVEVEHL 120
OY 121 ISQCHYOHGHINSYLPMLQDRDITLPAKGLDHIENILSYLDAKSLCAAEIVCKEWM 180
Db 121 ISQCHYOHGHINSYLPMLQDRDITLPAKGLDHIENILSYLDAKSLCAAEIVCKEWM 180
OY 181 RYVSDGMLMKKLLERWRTDLSLWGLAERKMGQYLFKNPPGNAFPNSFYALYPKII 240
Db 181 RYVSDGMLMKKLLERWRTDLSLWGLAERKMGQYLFKNPPGNAFPNSFYALYPKII 240
OY 241 ODITISNMRCGRHSIQRHCHRETSKGYCLOYDDOKIVSGLRDNTIKIMDKSTLECK 300
Db 241 ODITISNMRCGRHSIQRHCHRETSKGYCLOYDDOKIVSGLRDNTIKIMDKSTLECK 300
OY 301 RILGHTGSVLCLOYDERVYITGSSDSTVRVWVNTGEMNTLIHCEAVLHLPFNNGM 360
Db 301 RILGHTGSVLCLOYDERVYITGSSDSTVRVWVNTGEMNTLIHCEAVLHLPFNNGM 360
OY 361 VTCKSDRSIAVMDASPTDITLRVLYGHRAAVNVDFDKYIVSASGDTIKYMNSTC 420
Db 361 VTCKSDRSIAVMDASPTDITLRVLYGHRAAVNVDFDKYIVSASGDTIKYMNSTC 420
OY 421 EFVFTLNGHKGIGIACLOYRDLVYVSSGSDNTIRLMDIEGACLVLEGHELVRCIRFDN 480
Db 421 EFVFTLNGHKGIGIACLOYRDLVYVSSGSDNTIRLMDIEGACLVLEGHELVRCIRFDN 480
OY 481 KRIVSGAYDGKIKVMDLVAAIDPPAPAGTCLRTLVHSGRVFLQDFEPOIVSSSHDT 540
Db 481 KRIVSGAYDGKIKVMDLVAAIDPPAPAGTCLRTLVHSGRVFLQDFEPOIVSSSHDT 540
OY 541 ILIMDFLNDPAAQAEPSPSRITYTISR 569
Db 541 ILIMDFLNDPAAHAEPSPSRITYTISR 569

RESULT 12
AAV78584
ID AAV78584 standard; Protein: 632 AA.
AC
XX AAV78584;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human protein SEQ ID NO 1246.

```

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation.  
 OS Homo sapiens.  
 XX WO200157190-A2.  
 XX PD 09-AUG-2001.  
 XX PF 05-FEB-2001; 2001WO-US04098.  
 XX PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 PR 20-JUN-2000; 2000US-0598075.  
 PR 19-JUL-2000; 2000US-0620325.  
 PR 01-SEP-2000; 2000US-0654936.  
 PR 15-SEP-2000; 2000US-0663561.  
 PR 20-OCT-2000; 2000US-0693325.  
 PR 30-NOV-2000; 2000US-0728422.  
 XX (HYSE-) HYSEQ INC.  
 PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 XX WPI: 2001-476283/51.  
 DR N-PSDB; AAK51717.  
 XX Nucleic acids encoding polypeptides with cytokine-like activities,  
 PT useful in diagnosis and gene therapy -  
 PS Claim 20; Page 3505-3507; 6221pp; English.  
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAH80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAH80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.  
 XX SQ Sequence 632 AA;  
 Query Match 98.6%; Score 2992.5; DR 22; Length 632;  
 Best Local Similarity 90.0%; Pred. No. 1.9e-282;  
 Matches 569; Conservative 0; Mismatches 0; Indels 63; Gaps 1;  
 QY 1 MDPAAVIOEKALK-----14  
 DB 1 MDPAAVIOEKALKFMMFPRSWCPGMNTWARSRLWATSTRVQCSPMSLWIGCSLADS 60  
 QY 15 -----PANSSEDCNNGEPKRIIPKNSLKOTYNSCARLCLNOETV 57  
 DB 61 MSLRLKLYNGTALTALEFNSSEDCNNGEPKRIIPKNSLKOTYNSCARLCLNOETV 120  
 QY 58 CLASTAMKTEVCVAKTKLANGTSMIVPORKLSASYEKELCVKYEOMSESQOVAFV 117  
 DB 121 CLASTAMKTEVCVAKTKLANGTSMIVPORKLSASYEKELCVKYEOMSESQOVAFV 180  
 QY 118 EHLISOMCHYGHINSYIKPMLQDFETALPARGLDIAENILSYIDAKSICAAELVCK 177  
 DB 181 EHLISOMCHYGHINSYIKPMLQDFETALPARGLDIAENILSYIDAKSICAAELVCK 240

QY 178 EWTYVTSOGLMKLIERVWTDLSMRGLAERRGNGOYLFRKKPPDGNAPPNSFYRALYP 237  
 DB 241 EWTYVTSOGLMKLIERVWTDLSMRGLAERRGNGOYLFRKKPPDGNAPPNSFYRALYP 300  
 QY 238 KIIODIETIESNMRCGRHSLOIRHCRSETSKGVYCLQYDDOKIVSGLDNTIKTDKTL 297  
 DB 301 KIIODIETIESNMRCGRHSLOIRHCRSETSKGVYCLQYDDOKIVSGLDNTIKTDKTL 360  
 QY 298 ECKRILTGHTGSVLCQYDERVILITGSSDSYRVWDMVTGEMLNTLIHCEAVILRFENN 357  
 DB 361 ECKRILTGHTGSVLCQYDERVILITGSSDSYRVWDMVTGEMLNTLIHCEAVILRFENN 420  
 QY 358 GMAVTCSDKRSIAVWDMASPDITLIRVLYGHRAVNVVDDKIVYASGDRITKYWNT 417  
 DB 421 GMAVTCSDKRSIAVWDMASPDITLIRVLYGHRAVNVVDDKIVYASGDRITKYWNT 480  
 QY 418 STCEVFTLNGHRRGIACQYRDLRVYSGSSDNTIRLMDICGACRLVLEHHEELVNCIR 477  
 DB 481 STCEVFTLNGHRRGIACQYRDLRVYSGSSDNTIRLMDICGACRLVLEHHEELVNCIR 540  
 QY 478 FDNKRIVSGAYDGKIKVMDVLAALDPAPAGTLCRLTVEHSGVFRLOFDEQIVSSSH 537  
 DB 541 FDNKRIVSGAYDGKIKVMDVLAALDPAPAGTLCRLTVEHSGVFRLOFDEQIVSSSH 600  
 QY 538 DDTLIMFDLNDPAAQAEPPSPSKTYTISR 569  
 DB 601 DDTLIMFDLNDPAAQAEPPSPSKTYTISR 632  
 RESULT 13  
 AAM79566  
 ID AAM79566 standard; Protein: 654 AA.  
 AC AAM79566;  
 DT 06-NOV-2001 (first entry)  
 DE Human protein SEQ ID NO 3212.  
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation.  
 OS Homo sapiens.  
 XX WO200157190-A2.  
 XX PD 09-AUG-2001.  
 XX PF 05-FEB-2001; 2001WO-US04098.  
 XX PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 PR 20-JUN-2000; 2000US-0598075.  
 PR 19-JUL-2000; 2000US-0620325.  
 PR 01-SEP-2000; 2000US-0654936.  
 PR 15-SEP-2000; 2000US-0663561.  
 PR 20-OCT-2000; 2000US-0693325.  
 PR 30-NOV-2000; 2000US-0728422.  
 XX (HYSE-) HYSEQ INC.  
 PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 XX WPI: 2001-476283/51.  
 DR N-PSDB; AAK52699.  
 PT Nucleic acids encoding polypeptides with cytokine-like activities,  
 PT useful in diagnosis and gene therapy -



QY 118 EHLISOMCHYOHGHIINSYLKPMLOQDFTTALPANGLDHIAENILSYDAKSLCAAEIYCK 177  
 DB 203 EHLISOMCHYOHGHIINSYLKPMLOQDFTTALPANGLDHIAENILSYDAKSLCAAEIYCK 262  
 QY 178 EMYRTSDGMLMKKLIEMVRTDLSLWGLAERGMGOYLFPKNPPDGNAPNSFYRALYP 237  
 DB 263 EMYRTSDGMLMKKLIEMVRTDLSLWGLAERGMGOYLFPKNPPDGNAPNSFYRALYP 322  
 QY 238 KIIDITETISNMFCGRHSIQRHCRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTL 297  
 DB 323 KIIDITETISNMFCGRHSIQRHCRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTL 382  
 QY 298 ECKRLTGHGTSVCLQYDERVITITGSSDSTVRYWVDNTGEMNTLHHCEAVLHLRFNN 357  
 DB 383 ECKRLTGHGTSVCLQYDERVITITGSSDSTVRYWVDNTGEMNTLHHCEAVLHLRFNN 442  
 QY 358 GMMVTCSDRSIAVWMAASPTDITLRRVLYGHRAAVNVVDFDDKYIVSAGDRTIKVWNT 417  
 DB 443 GMMVTCSDRSIAVWMAASPTDITLRRVLYGHRAAVNVVDFDDKYIVSAGDRTIKVWNT 502  
 QY 418 STCEFYRTLNGHKGKGLACIQRDLVYVSGSSDNTIRLMDIECGACLRVLEGEHELYRCIR 477  
 DB 503 STCEFYRTLNGHKGKGLACIQRDLVYVSGSSDNTIRLMDIECGACLRVLEGEHELYRCIR 562  
 QY 478 FDNKRIYSGAYDGKIKVWDVLAALDPAPAGTLCRLTIVHSGRVPRLOPDEFQIYSSSH 537  
 DB 563 FDNKRIYSGAYDGKIKVWDVLAALDPAPAGTLCRLTIVHSGRVPRLOPDEFQIYSSSH 622  
 QY 538 DDTLLIMDFLNDPAQAQEPSPSRITYTISR 569  
 DB 623 DDTLLIMDFLNDPAQAQEPSPSRITYTISR 654

## RESULT 15

AAW79568  
 ID AAW79568 standard; Protein; 654 AA.

AC AAW79568;

DT 06-NOV-2001 (first entry)

DE Human protein SEQ ID NO 3214.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorder; arthritis; inflammation.

OS Homo sapiens.

PN WO200157190-A2.

XX 09-AUG-2001.

PD 05-FEB-2001; 2001WO-US04098.

PF 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.

PR 01-SEP-2000; 2000US-0654936.

PR 15-SEP-2000; 2000US-0663561.

PR 20-OCT-2000; 2000US-0693325.

PR 30-NOV-2000; 2000US-0728422.

XX (HYTE-) HYTEQ INC.

XX PA

XX PI Tang YF, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

XX PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX DR WPI; 2001-476283/51.

DR N-PSDB; AAK52701.

XX Nucleic acids encoding polypeptides with cytokine-like activities,

PT useful in diagnosis and gene therapy -

XX Claim 20; Page 286-287; 6221pp; English.

CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAW78323-AAW80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAW80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.

CC Sequence 654 AA:

Query Match 98.5%; Score 2989.5; DR 22; Length 654;  
 Best Local Similarity 89.9%; Pred. No. 3.9e-282;  
 Matches 568; Conservative 1; Mismatches 0; Indels 63; Gaps 1;

QY 1 MDPAAVLAQERKLEKPM----- 16

DB 23 MDPAAVLAQERKLEKPMFERSMCPQWNTMARSLVATSTSVGCSMPRSLMGCCSLADS 82

QY 17 -----NSEREDCNNGEPKRIIPEKNSLQOTNSCARCLDNETV 57

DB 83 MPELRCLYNPGTGAITAFQNSEREDCNNGEPKRIIPEKNSLQOTNSCARCLDNETV 142

QY 58 CLASTAMKTENCVAATKLANGTSSMIVPKORKLSASYEKKEKELCVKFFQWSESDQYEFV 117

DB 143 CLASTAMKTENCVAATKLANGTSSMIVPKORKLSASYEKKEKELCVKFFQWSESDQYEFV 202

QY 118 EHLISOMCHYOHGHIINSYLKPMLOQDFTTALPANGLDHIAENILSYDAKSLCAAEIYCK 177

DB 203 EHLISOMCHYOHGHIINSYLKPMLOQDFTTALPANGLDHIAENILSYDAKSLCAAEIYCK 262

QY 178 EMYRTSDGMLMKKLIEMVRTDLSLWGLAERGMGOYLFPKNPPDGNAPNSFYRALYP 237

DB 263 EMYRTSDGMLMKKLIEMVRTDLSLWGLAERGMGOYLFPKNPPDGNAPNSFYRALYP 322

QY 238 KIIDITETISNMFCGRHSIQRHCRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTL 297

DB 323 KIIDITETISNMFCGRHSIQRHCRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTL 382

QY 298 ECKRLTGHGTSVCLQYDERVITITGSSDSTVRYWVDNTGEMNTLHHCEAVLHLRFNN 357

DB 383 ECKRLTGHGTSVCLQYDERVITITGSSDSTVRYWVDNTGEMNTLHHCEAVLHLRFNN 442

QY 358 GMMVTCSDRSIAVWMAASPTDITLRRVLYGHRAAVNVVDFDDKYIVSAGDRTIKVWNT 417

DB 443 GMMVTCSDRSIAVWMAASPTDITLRRVLYGHRAAVNVVDFDDKYIVSAGDRTIKVWNT 502

QY 418 STCEFYRTLNGHKGKGLACIQRDLVYVSGSSDNTIRLMDIECGACLRVLEGEHELYRCIR 477

DB 503 STCEFYRTLNGHKGKGLACIQRDLVYVSGSSDNTIRLMDIECGACLRVLEGEHELYRCIR 562

QY 478 FDNKRIYSGAYDGKIKVWDVLAALDPAPAGTLCRLTIVHSGRVPRLOPDEFQIYSSSH 537

DB 563 FDNKRIYSGAYDGKIKVWDVLAALDPAPAGTLCRLTIVHSGRVPRLOPDEFQIYSSSH 622

QY 538 DDTLLIMDFLNDPAQAQEPSPSRITYTISR 569

DB 623 DDTLLIMDFLNDPAQAQEPSPSRITYTISR 654

Mon Feb 24 10:42:30 2003

us-09-601-168b-2.rag

Page 13

Search completed: February 20, 2003, 09:53:46  
Job time : 89 secs

...

...

...

...

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 09:53:52 ; Search time 14 seconds  
(without alignments)

1685.716 Million cell updates/sec

Title: US-09-601-168b-2

Perfect score: 3034  
Sequence: 1 MDPAAVLAQKALKFMSSE.....PAAQAPPPSPSRITYTISR 569

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3006	99.1	605	FW1A_HUMAN	Q9Y297 homo sapien
2	2597	85.6	518	TRCB_XENLA	Q91854 xenopus lae
3	2384.5	78.6	542	FW1B_HUMAN	Q9UBD1 homo sapien
4	1638.5	54.0	665	L123_CAEEL	Q09990 caenorhabd
5	690	22.7	506	POF8_SCHPO	Q09855 schizosacch
6	590.5	19.5	605	POF1_SCHPO	P87053 schizosacch
7	575	19.0	678	SCOB_EMENTI	Q00659 emeticella
8	545	18.0	640	MT30_YEAST	P39014 saccharomyc
9	531.5	17.5	650	SCO2_YEAST	Q01277 neurospora
10	520	17.1	579	SE10_CAEEL	Q93794 caenorhabd
11	477.5	15.7	684	CC4_CANAL	P53699 candida alb
12	455.5	15.0	1356	HET1_PODAN	Q00808 podospora a
13	453	14.9	775	POF1_SCHPO	P87060 schizosacch
14	399	13.2	779	CC4_YEAST	P07834 saccharomyc
15	396.5	13.1	703	POP2_SCHPO	Q01470 schizosacch
16	392	12.9	732	KMBH_DICDI	P90648 dictyostell
17	374	12.3	1526	YV46_ANASP	Q8Y112 anabaena sp
18	373.5	12.3	1258	YS00_ANASP	Q8Y112 anabaena sp
19	373.5	12.3	1683	YL24_ANASP	Q8Y112 anabaena sp
20	354	11.7	409	L1S1_HUMAN	P43034 homo sapien
21	354	11.7	409	L1S1_MOUSE	P43035 mus musculu
22	353	11.6	409	L1S1_BOVIN	P43033 bos taurus
23	341	11.2	515	YCM2_YEAST	P25382 saccharomyc
24	337	11.0	1146	KMHA_DICDI	D42527 dictyostell
25	334.5	11.0	422	FBW2_HUMAN	Q9UKT8 homo sapien
26	325.5	10.7	361	WDS_DROME	Q9V318 drosophila
27	324.5	10.7	361	WDR3_HUMAN	Q9UG99 homo sapien
28	318	10.5	422	FBW2_MOUSE	Q60584 mus musculu
29	313.5	10.3	376	YK14_CAEEL	Q17963 caenorhabd
30	312.5	10.3	714	YUL2_YEAST	P47025 saccharomyc
31	307.5	10.1	742	PKMA_THRCU	P49695 thermomonos
32	307.5	10.1	1693	Y163_STYX3	O5553 synecocyst
33	306	10.1	704	T2D4_DROME	P49846 drosophila

34	298.5	9.8	800	1	T2D4_HUMAN	Q1542 homo sapien
35	294.5	9.7	1249	1	APAF_RAT	Q9EPV5 rattus norv
36	292	9.6	659	1	YK16_YEAST	P36130 saccharomyc
37	290.5	9.6	614	1	TU11_SCHPO	Q09715 schizosacch
38	289	9.5	1249	1	APAF_MOUSE	Q88879 mus musculu
39	287.5	9.5	327	1	GBLP_BRANA	Q39336 brassica na
40	287	9.5	1248	1	APAF_HUMAN	Q14727 homo sapien
41	286.5	9.4	327	1	GBLP_ARATH	Q24456 arabidopsis
42	285.5	9.4	798	1	T2D4_YEAST	P38129 saccharomyc
43	283	9.3	473	1	PRP5_SCHPO	Q13615 schizosacch
44	283	9.3	713	1	TUPL1_YEAST	P16649 saccharomyc
45	281.5	9.3	682	1	TUPL1_KL01A	P56094 kluyveromyc

## ALIGNMENTS

RESULT 1	ID	FW1A_HUMAN	STANDARD:	PRT:	605 AA.
AC	Q9Y297	Q9Y213:			
DT	16-OCT-2001	(Rel. 40, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	F-box/WD-repeat protein 1B (F-box and WD-repeats protein beta-TrCP)				
DE	(E3RSIXAPPAB) (pikappabalpha-E3 receptor subunit).				
GN	FBXW1A OR FBW1A OR BTRCP OR BTRCP				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RX	MEDLINE=99075339; PubMed=9859996;				
RA	Varon A., Hatzubai A., Davis M., Layon I., Amit S., Manning A.M.,				
RA	Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.,				
RT	"Identification of the receptor component of the I-kappa-alpha-				
RL	ubiquitin ligase.";				
RL	Nature 396:590-594(1998).				
RN	[2]				
RP	SEQUENCE FROM N.A. (ISOFORM 2).				
RC	TISSUE=Lymphoid;				
RX	MEDLINE=98325370; PubMed=9660940;				
RA	Maroltin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,				
RA	Thomas D., Strebel K., Benarous R.;				
RT	"A novel human WD protein, h-beta TrCP, that interacts with HIV-1 Ypu				
RT	connects Cdk4 to the ER degradation pathway through an F-box motif.";				
RL	Mol. Cell 1:565-574(1998).				
RN	[3]				
RP	SEQUENCE FROM N.A. (ISOFORM 2).				
RX	MEDLINE=20003060; PubMed=10531035;				
RA	Cenciarrelli C., Chaur D.S., Guardavaccaro D., Parks W., Vidal M.,				
RA	Pagano M.;				
RT	"Identification of a family of human F-box proteins.";				
RL	Curr. Biol. 9:1177-1179(1999).				
RN	[4]				
RP	CHARACTERIZATION.				
RX	MEDLINE=99145464; PubMed=9990852;				
RA	Winston J.T., Strack P., Beer-Romero P., Chu C.Y., Elledge S.J.,				
RA	Harper J.W.,				
RT	"The SCF(beta-TRCP)-ubiquitin ligase complex associates specifically				
RT	with phosphorylated destruction motifs in I-kappa-B-alpha and				
RT	beta-catenin and stimulates I-kappa-B-alpha ubiquitination in vitro.";				
RL	Genes Dev. 13:270-283(1999).				
CC	- FUNCTION: BINDS SPECIFICALLY TO PHOSPHORYLATED IKB ALPHA				
CC	(PIKAPPABALPHA) AND PHOSPHORYLATED BETA-CATENIN AND PROMOTES THEIR				
CC	UBIQUITINATION AND DEGRADATION.				
CC	- SUBUNIT: PART OF A SCF (SKP1-CULIN-F-BOX) PROTEIN LIGASE COMPLEX.				
CC	- SUBCELLULAR LOCATION: Cytoplasmic.				
CC	- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2; ARE				
CC	PRODUCED BY ALTERNATIVE SPLICING.				
CC	- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.				
CC	- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).				

CC This SWISS-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF101784; AAD08702.1; -  
 DR EMBL: Y14153; CAA7572.1; -  
 DR EMBL: AF129530; AAF04464.1; -  
 DR Genew: HGNC:1144; BTRC.  
 DR MIM: 603482; -  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00320; WD40; 7.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; F-box; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 6.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 DR Uni conjugation pathway; Repeat; WD repeat; Alternative splicing.  
 KM DOMAIN 190 228 F-box.  
 FT REPEAT 301 338 WD 1.  
 FT REPEAT 341 378 WD 2.  
 FT REPEAT 381 418 WD 3.  
 FT REPEAT 424 461 WD 4.  
 FT REPEAT 464 503 WD 5.  
 FT REPEAT 505 541 WD 6.  
 FT REPEAT 553 590 WD 7.  
 FT REPEAT 590 605 WD 7.  
 FT VARSPLIC 17 52 MISSING (IN ISOFORM 2).  
 SQ SEQUENCE 605 AA; 66866 MM; 4C67F3B7E400FD37 CRC64;

Query Match 99.1%; Score 3006; DB 1; Length 605;  
 Best Local Similarity 94.0%; Pred. No. 7.7e-219;  
 Matches 569; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 MDPAEAVLOKALKFM-----NSSREDC 24  
 DB 1 MDPAEAVLOKALKFM-----NSSREDC 60  
 QY 25 NNGEPPKIIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTENCYAKTKLANGTSMIV 84  
 DB 61 NNGEPPKIIPEKNSLRQTYNSCARLCLNOETVCLASTAMKTENCYAKTKLANGTSMIV 120  
 QY 85 PKORKLASATEKEKELCVKFEQWSSDOVEFEVHLISQCHYQGHINSYKPLMRDQF 144  
 DB 121 PKORKLASATEKEKELCVKFEQWSSDOVEFEVHLISQCHYQGHINSYKPLMRDQF 180  
 QY 145 ITPALPARGDHTIENILSYDAKSLCAAEIYCKEYRVSDDGLMKKILERRWRDLSLR 204  
 DB 181 ITPALPARGDHTIENILSYDAKSLCAAEIYCKEYRVSDDGLMKKILERRWRDLSLR 240  
 QY 205 GLAERRGMOQYLFKNRPPDGNAPPNSEFYALYPKIIQDIETTESNMRGRHSIORIHCS 264  
 DB 241 GLAERRGMOQYLFKNRPPDGNAPPNSEFYALYPKIIQDIETTESNMRGRHSIORIHCS 300  
 QY 265 ETSKGYTCLOYDQKIVSGIRDTIKIMDKNLTICKRRIITGHTGSVLCLOYDERVITGS 324  
 DB 301 ETSKGYTCLOYDQKIVSGIRDTIKIMDKNLTICKRRIITGHTGSVLCLOYDERVITGS 360  
 QY 325 SDSTVAVMDVNTGEMLNTLIHCEAVLHLRFNNGMMVTCSSKRSIAVMDASPTDITLRR 384  
 DB 361 SDSTVAVMDVNTGEMLNTLIHCEAVLHLRFNNGMMVTCSSKRSIAVMDASPTDITLRR 420  
 QY 385 VLVGHRAAVNVVDFDDKTYIVSASGDFTIKVMNTSTCEFTVINGHGRGTAICLOYDRILVV 444  
 DB 421 VLVGHRAAVNVVDFDDKTYIVSASGDFTIKVMNTSTCEFTVINGHGRGTAICLOYDRILVV 480

QY 445 SGGSDTIRLMDIEGACALRVIEGHELVRCIRPNKRIVSGAYDGKTKVMDVLAALDPR 504  
 DB 481 SGGSDTIRLMDIEGACALRVIEGHELVRCIRPNKRIVSGAYDGKTKVMDVLAALDPR 540  
 QY 505 AAGTCTCLTVLHESGRVRLQDFEPOVSSSHDPTILMPILNPPAQAAPPSRSTY 564  
 DB 541 AAGTCTCLTVLHESGRVRLQDFEPOVSSSHDPTILMPILNPPAQAAPPSRSTY 600

QY 565 TYISR 569  
 DB 601 TYISR 605

RESULT 2  
 TRCB\_XENLA STANDARD; PRT; 518 AA.  
 ID 091854; P70037; P70038;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Beta-Trip (Beta-transducin repeat-containing protein).  
 GN FBXW1 OR BTRCP.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;  
 OC Xenopodidae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=93330289; PubMed=8393141;  
 RX Spevak W., Keiper B.D., Stratowa C., Castanon M.J.;  
 RT "Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in  
 RT anaphase are rescued by Xenopus CDNs encoding N-ras or a protein  
 RT with beta-transducin repeats.";  
 RL Mol. Cell. Biol. 13:4953-4966(1993).  
 RN [2]  
 RP SEQUENCE OF 302-518 FROM N.A.  
 RA MEDLINE=97109804; PubMed=8952061;  
 RX Hudson J.W., Alarcon V.B., Elinson R.P.;  
 RT "Identification of new localized RNAs in the Xenopus oocyte by  
 RT differential display PCR.";  
 RL Dev. Genet. 19:190-198(1996).  
 CC - FUNCTION: Probably recognizes and binds to some phosphorylated  
 CC proteins and promotes their ubiquitination and degradation.  
 CC - SUBUNIT: PART OF A SCF (SKP1-CULIN-F-BOX) PROTEIN LIGASE COMPLEX  
 CC (BY SIMILARITY).  
 CC - DEVELOPMENTAL STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-  
 CC MAURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC  
 CC GENE TRANSCRIPTION BEGINS FOLLOWING THE MIDBLASTULA TRANSITION. DO  
 CC NOT INCREASE IN ABUNDANCE IN THE GASTRULA, NEURULA, TAILBUD, OR  
 CC TADPOLE EMBryo.  
 CC - SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC - SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC - SIMILARITY: STRONG, TO C.ELEGANS K10B2.1.

CC This SWISS-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M98268; AAA02810.1; -  
 DR EMBL: U63921; AAB49671.1; -  
 DR EMBL: U63922; AAB49672.1; -  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR ProDom: PD000018; WD40; 4.





DB 203 NNNKRNKMTIRINCSSEKSGYCCQYDDDKIVSGLRDNTIKIMDKDYSRLLSGHT 262  
 QY 308 GSVLCQYDERVITITGSSSTVAVDVTGEMNLTHHCCEVLIHFRFNNMMVTCSSKR 367  
 DB 263 GSVLCQYDERVITITGSSSTVAVDVTGEMNLTHHCCEVLIHFRFNNMMVTCSSKR 322  
 QY 368 STAVMDASPTDITLRLVGVHRAAVNVVDKDYIVSASGRTIKVMTSTCEVYRLN 427  
 DB 323 STAVMDASPTDITLRLVGVHRAAVNVVDKDYIVSASGRTIKVMTSTCEVYRLN 382  
 QY 428 GHRGACIQLQYRDLRVYSSSDMTIRLMDIEGACILRVIEGHELVRCIRFENKRVISA 487  
 DB 383 GHRGACIQLQYRDLRVYSSSDMTIRLMDIEGACILRVIEGHELVRCIRFENKRVISA 442  
 QY 488 YDKIKVMDVIALDPRAPAGTCLRTLVHSGRPFRLQDEFQYSSSHDDITILMDL 547  
 DB 443 YDKIKVMDVIALDPRAPAGTCLRTLVHSGRPFRLQDEFQYSSSHDDITILMDL 502  
 QY 548 NDPAQAEPSPSPRT 563  
 DB 503 DAP-----PGLSPST 513

RESULT 5  
 POFB\_SCHPO STANDARD; PRT; 506 AA.  
 ID 009855: 09P7V1:  
 AC 01-FEB-1996 (Rel. 33, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE F-box/WD-repeat protein pof1.  
 GN POF1 OR SPAC296.01 OR SPAC30.05.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 ON NCBI\_TaxID=4896;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA Harrison C.L., Toda T.:  
 RT "Systematic genome-wide analysis of F-box protein-encoding genes in  
 RL fission yeast."  
 RN Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 RC [2]  
 RP SEQUENCE FROM N.A.  
 RC SRRATN-972:  
 RX MEDLINE-21848401; PubMed-11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,  
 RA Welteens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hlibert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eber P., Zimmermann W., Medler H., Mambutti R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Kochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Sipkavskii G., Ussery D., Barrell B.G., Nurse P.:  
 RT "The genome sequence of Schizosaccharomyces pombe."

RL Nature 415:871-880(2002).  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AB061694; BAB5543.1;  
 DR EMBL: AL136538; CAB66464.1;  
 DR EMBL: Z66525; CA91423.1;  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR ProDom: PD000018; WD40; 3.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 8.  
 DR PROSITE: PS00181; FBOX; 1.  
 DR PROSITE: PS00676; WD\_REPEATS\_1; 3.  
 DR PROSITE: PS00082; WD\_REPEATS\_2; 6.  
 DR PROSITE: PS0294; WD\_REPEATS\_REGION; 1.  
 KW Repeat: WD repeat.  
 FT DOMAIN 70 116 F-BOX.  
 FT REPEAT 219 256 WD 1.  
 FT REPEAT 259 298 WD 2.  
 FT REPEAT 301 338 WD 3.  
 FT REPEAT 345 386 WD 4.  
 FT REPEAT 388 426 WD 5.  
 FT REPEAT 427 464 WD 6.  
 FT REPEAT 468 505 WD 7.  
 SQ SEQUENCE 506 AA; 58257 MW; CEF34D4EFPBC2E10 CRC64;

Query Match 22.7%; Score 690; DB 1; Length 506;  
 Best Local Similarity 30.4%; Pred. No. 1.4e-44;  
 Matches 163; Conservative 104; Mismatches 167; Indels 102; Gaps 15;

QY 67 ENCVAK-----TLANGTSMIVPKOKISASTYKEKELCYKFFEQWSSDOVEPEHLIS 122  
 DB 8 KNVSKVSDLTSCSDSFSTSSPVPCLNPLS-----HNNRIDLRLDLA 50  
 QY 123 OKCHYGHINSYLRKMLORDFTLAPAGLDHIAENITSLYDAKSLCAELCKEYRY 182  
 DB 51 SLKSGEVAVVYNNVSLTDFTEVFP-----EVLRFVSYLDOLDICKKILMSKRMRL 106  
 QY 183 TSDGMLMKRLI-----ERNVRDLSMRG-----LAERGWG----- 213  
 DB 107 LEDPGIKALYMQKGFVENVLEFEAWRTRHKFPQPFENFLKQONIGYGTMLPQ 166  
 QY 214 QYLFKKPPDGNAPPSFRALYPKLIDILETIESMRCGRSLDRIHRS----- 264  
 DB 167 QYLF-----DSNGRPLNMSYLY-----KEHALDSMWRGRPLVSTFNPSIRFPADQDF 217  
 QY 265 -ETSKGVYCLQDDQKIVGLRDNITIKIDKNTLLECKRLITGHTSVLCLOYDER-VII 321  
 DB 218 RALDSYVCEVDDDELMVSGSKDRIVSYVDVSRFLYLYLGHSSSVLCIDPCRRLNLY 277  
 QY 322 TGGSSDSTVRVMDVNTGEMNLTHHCCEVLIHFRFNNMMVTCSSKRSLAVW--DMASPTD 379  
 DB 278 SGSSDSTIIMQNNRPLKVFEGHTDNLGVVSENYSISSRHTAVWRDLATSPA 337  
 QY 380 ITLRRLVGVHRAAVNVVDKDK--IVSASGDRITKVMNTSTCEVYRLNGHKGRIACQ 437  
 DB 338 ACN-HVLRGHLASVNSVOSSKGLIVTASDDTITFTWITTGHCIRITHAQRGIAQ 396  
 QY 438 YRDLRVYSSSDMTIRLMDIEGACILRVIEGHELVRCIRFENKRVISGAYDGKTKVMDL 497  
 DB 397 YNCKFLVSGSSDLTIRIFASSGKILRLMLOGHDDLIRTVRFNDKRVISGAYGYRINN- 455

QY 498 VAAADPRAPAGTLCRTLVH-----SGRVFLQDFEIOYSSHDDTLLWDF 546  
 Db 456 -----FNTGECHCVLANSRNSKVFGLQDFHRIIACHTSSHTIWNF 497

RESULT 6  
 ID POP1\_SCHPO STANDARD; PRT; 605 AA.  
 AC P87053;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE F-box/MD-repeat protein pop1 (SKP1-binding protein 1).  
 GN POP1 OR SAK1 OR SPAC57A10.05c.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_Taxid=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA Katayama S., Kitamura K., Toda T.;  
 RT "Systematic genome-wide analysis of F-box protein-encoding genes in  
 RT fission yeast."  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Soutos J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., O'Neill C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Voicakert G., Aert R., Robben J., Grymoprez B.,  
 RA Meltens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Beyer P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Delaure V., Motier S.,  
 RA Galibert F., Aves S.D., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerretti L., Lowe T., McCombie M.R., Paulsen I., Potlashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe".  
 RL Nature 415:871-880(2002).  
 CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated  
 CC proteins and promotes their ubiquitination and degradation.  
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CUL1IN-F-BOX) PROTEIN LIGASE COMPLEX  
 CC (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: AB032410; BAA84528.1; -

DR EMBL: 294864; CAB08168.1; -  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR PRODOM: PD000018; WD40; 5.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD REPEATS\_1; 2.  
 DR PROSITE: PS50082; WD REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD REPEATS\_REGION; 1.  
 KW Ubl conjugation; Repeat; WD repeat.  
 FT DOMAIN 107 153  
 FT REPEAT 271 299 WD 1.  
 FT REPEAT 311 339 WD 2.  
 FT REPEAT 350 379 WD 3.  
 FT REPEAT 390 420 WD 4.  
 FT REPEAT 432 460 WD 5.  
 FT REPEAT 472 500 WD 6.  
 FT REPEAT 510 538 WD 7.  
 SQ SEQUENCE 605 AA; 67110 MW; 7118C9379EC5C1F0 CRC64;

Query Match 19.5%; Score 590.5; DB 1; Length 605;  
 Best Local Similarity 29.0%; Freq. No. 5.5e-37;  
 Matches 156; Conservative 86; Mismatches 189; Indels 107; Gaps 12;

QY 80 SSMIVPKRKILASAYEKKEKELCKYFEQWSE---SDQVEFVHILISQCHVQHGHINSYL 136  
 Db 46 SSM-----HNEISGLSEKSNQREAVMAFSEASCSERIALOGILINNCSSLSFASSTL 101  
 QY 137 KMLQDRFTLALPANGDLHIAENILSYDAKSLCAELVCKEWTYVSDGMWKILERM 196  
 Db 102 DSLVRLDFSLTPV---EISFRILSFLDARSICQAAVSKHWEKELADDDVIMRMCEQH 157  
 QY 197 VTTDSLWGLMERRGWGYLFKN----- 219  
 Db 158 INRK-----CEKCGKGLPLERNTLYAKASIQKRYERLTKRGVDQAHSSPYKAKLD 211  
 QY 220 -----KPDGNAPPNSE-----YRALPKTIQDIETIESNMRCGRSLQ 258  
 Db 212 DYPTSSNETTISVYPPSPNSDSKFFLPFKTRPMKEVYAEKRC---VECWNRGR----- 263  
 QY 259 RHCH-----SEPSGYVQLODDQDIYVGLNDNTIKMDKTLCKRLVGHSTVYCLQ 314  
 Db 264 ---CQVYLSGHSDVWCLQVRLNLSGSDYDATTILMNLTATFOVALLEGHSSGYVCLQ 320  
 QY 315 YDERVITNGSSDSYVRWVDVTGEMLTLLIHCEAVLHIFRNQGMATYCSKDRSIAYWDM 374  
 Db 321 FDCCKLISGMDKTRIMRNVTSEGISLHGHDSVYCLTFPDSLLVSGSADCTVAKIMHF 380  
 QY 375 ASPDITLRLVGVHRAAVNYDF--DDKIVYASGSDRTIKVWNTSVCEVFTLNGHKRG 432  
 Db 381 SGGRKTRLR---GHTGPNVSVRIIRDRGLVLSGSDSTIKWISLETNLCIHTPSAHICP 436  
 QY 433 IACLOVRDLVVGSSSDNTIRLMDIEGACLRVLEGEHELRCIRPNKRIYSAAYGKI 492  
 Db 437 VQSLADLSRFLSCSLDCTIKQMDIEKKKCVHTLFGHIGCWELIADHLRIASAHGIV 496  
 QY 493 KVMDLVAAIDPRAPAGTLCRTLVHSGRVFLQDFEIOYSSHDDTLLWDFLNDP 550  
 Db 497 KVMKECE-----CVHTLKNHSEPTVSVALGCEVVGSGSDCKTYLWLFNNAP 543

RESULT 7  
 ID SCOB\_EMENT STANDARD; PRT; 678 AA.  
 AC Q00659;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-OCT-2001 (Rel. 40, Last annotation update)  
 DE Sulfur metabolite repression control protein.

GN SCOMB OR MAPBL.  
 OS Emericella nidulans (Aspergillus nidulans).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 ON Eubacteria; Trichocomaceae; Emericella.  
 RX NCBI\_TaxId=5072;  
 RP SEQUENCE FROM N.A.  
 RA Natorff R.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DDJ databases.  
 CC -1- FUNCTION: REGULATORY PROTEIN THAT CONTROL SULFUR METABOLITE  
 REPRESSION.  
 CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE MET30/SCOMB/SCON-2 FAMILY OF WD-REPEAT  
 PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U21220; AAC15905.1; -  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR Pfam: PR00320; GPROTEINRPT.  
 DR PRINTS: PR00320; GPROTEINRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS00181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 4.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 7.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 DR KW Transcription regulation; Repeat; WD repeat.  
 FT DOMAIN 178 224 F-BOX.  
 FT REPEAT 347 375 WD 1.  
 FT REPEAT 387 415 WD 2.  
 FT REPEAT 427 455 WD 3.  
 FT REPEAT 466 496 WD 4.  
 FT REPEAT 508 543 WD 5.  
 FT REPEAT 553 595 WD 6.  
 FT REPEAT 607 635 WD 7.  
 FT REPEAT 647 675 WD 8.  
 FT SEQUENCE 678 AA; 76070 MW; D840D452E37B4C53 CRC64;  
 SO  
 Query Match 19.0%; Score 575; DB 1; Length 678;  
 Best Local Similarity 26.9%; Pred. No. 9,4e-36;  
 Matches 149; Conservative 83; Mismatches 168; Indels 154; Gaps 14;

QY 354 RENNOMATYCKSRISAIVADNASPTDITLRVYLGHRAAVV--VDDDKIYVASGDRT 411  
 DB 436 HEDASITLASSGVDTKTVKILWNEFKSTFLR-----GHTDWNAVAVRDISSRTVFSASDCT 491  
 QY 412 IYVWNTSTCEFYRTLNH-----KRDIAC----- 435  
 DB 492 VRLMDLDTKTCIRTFHGVGOVYVPLPRFEFEHDEACENDLSTTSGDANPPIQA 551  
 QY 436 -----LQYRD-----RLVYSGSDNTIRLMDIEGACLRVLEGHEEL 472  
 DB 552 SNGLEPNMAVSSOASAFGTSFNDGRAAPRYWTSALDSTIRLMETTRCRLRTFFGHLEG 611  
 QY 473 VQICRFENKRIYSGAVDGKIKYMDLVAALDPPAPAGTLCRLTVHSGRVRLQDFEQI 532  
 DB 612 VVALGADTLRIYSGAEDRMKIMD-----PRTGKCFRTFGHSGPVICIGIDSRF 662  
 QY 533 VSSSHDITLIMDF 546  
 DB 663 ATGSEDCVEMYSF 676  
 RESULT 8  
 MT30\_YEAST STANDARD: PRT: 640 AA.  
 AC P39014;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE MET30 protein.  
 GN MET30 OR YIL046W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 RX NCBI\_TaxId=4932;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=2180-1A;  
 RX MEDLINE=96069360; PubMed=8524217;  
 RA Thomas D., Kuras L., Barbey R., Charest H., Blaiseau P.L.,  
 RA Surdin-Kerjan Y.;  
 RT "Met30p, a yeast transcriptional inhibitor that responds to S-  
 adenosylmethionine, is an essential protein with WD40 repeats";  
 RL Mol. Cell. Biol. 15:6526-6534(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=5288C / AB972;  
 RA Churchill C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,  
 RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,  
 RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,  
 RA Rajadream M.A., Riles L., Rowley N., Skelton J., Smith V.,  
 RA Walsh S.V., Whitehead S.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DDJ databases.  
 CC -1- FUNCTION: NEGATIVELY REGULATES SULFUR AMINO ACIDS BIOSYNTHESIS  
 GENES EXPRESSION.  
 CC -1- SUBUNIT: SEEMS TO INTERACT WITH MET4.  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: BELONGS TO THE MET30/SCOMB/SCON-2 FAMILY OF WD-REPEAT  
 PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z46861; CA86905.1; -  
 DR EMBL: L26505; AAA96717.1; -  
 DR SGD: S0001308; MET30.  
 DR InterPro: IPR001810; F-box.

DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPOTEINRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 6.  
 DR PROSITE: PSS0181; FBOX; 1.  
 DR PROSITE: PSS0678; WD\_REPEATS\_1; 4.  
 DR PROSITE: PSS0082; WD\_REPEATS\_2; 6.  
 DR PROSITE: PSS0294; WD\_REPEATS\_REGION; 1.  
 DR Transcription regulation: Methionine biosynthesis;  
 KW Cysteine biosynthesis: Repeat; WD repeat.  
 FT DOMAIN 161 227  
 F-BOX.  
 FT REPEAT 300 328 WD 1.  
 FT REPEAT 340 368 WD 2.  
 FT REPEAT 380 408 WD 3.  
 FT REPEAT 419 449 WD 4.  
 FT REPEAT 461 499 WD 5.  
 FT REPEAT 509 538 WD 6.  
 FT REPEAT 550 578 WD 7.  
 FT REPEAT 607 635 WD 8 (POTENTIAL).  
 FT CONFLICT 61 61 M -> I (IN REF. 1).  
 SQ SQUONCE 640 AA; 72835 MW; 5135DABCAZE1EB97 CRC64;

Query Match 18.0%; Score 545; DB 1; Length 640;  
 Best Local Similarity 28.9%; Pred. No. 1.6e-33;  
 Matches 167; Conservative 85; Mismatches 202; Indels 124; Gaps 17;

QY 34 IPEKNSLRQTYNSCARLCLN-QETVCLASTAMKTENCVAK-TKLANGSTSMIVPKRKLS 91  
 DB 86 LPEYNEFKFCFYRHNPDIQSPHTACVQKDKRTQGINNNAIKLPQESDIIHIIISKS 145  
 QY 92 ASYEKKELCVKFEQMSSEDOVEFEHLISOMQHGHINSYKPMLODFITALPAR 151  
 DB 146 NSMDKTRKL-----ILDGILSTISCFPLSISLVYHMIKIPISILP- 188  
 QY 152 GLDHIENLISYDAKSLCAELVCKEYRVTSQGLMKLIERVVRIDSLKRLAER- 210  
 DB 189 -QELSLKILSYDCOSLONATRVCKKMKLADDDRVYHMEOH- 236  
 QY 211 -GMCQYLFKNK-----PPDGNAPNSFYALYPKIIQDIETTESNRCGRHSIO 258  
 DB 237 PNCQWGLPLIHKMRARIQONSTGSSSNADIQOTRPWKVIVREPKVESNMKRG- 291  
 QY 259 RIHCSETSK---GYCYQYDOKIVSGRLNTIKIMKNTLECKRILITGHTGSLCQ 314  
 DB 292 --HCRIQEKGMHGVLTIQENYRLFTGSDYSTIGIMDLFTGKLIRLSGSHDGVKTL 349  
 QY 315 YDERVITIGSSDSTYAVWVDVTGEMLNTLIHCEAVLHLRFNNGMMVYCSKRSIAVDM 374  
 DB 350 FDDRKLITGSLDKITIRVWNYITGECISTYRGHSYLSVDSQKVIYSGSADKIVKWHV 409  
 QY 375 ASPDTILIRVAVGRAVNVVDFDKYI--VSASGDRTIKYWNTSTGEFVFTLNGH- 429  
 DB 410 ESRCTYTLR---GHEWVNCVKLHPKPSFSCSDDTIRMDIRITNSCLKVFGVHGQ 465  
 QY 430 KRGIACLOYRD--RVVSGSS----- 448  
 DB 466 VKKIIPLIKIDENLATONTISDSSPPDDPTMTGACADSDTPSPNOETVLIDENIPYHL 525  
 QY 449 -----DNTIRLMDIEGACGLAVLIEGHELVRCIRPDNRRIYGAVDGKIWDVLAALDP 503  
 DB 536 LSCGIDNTIKIMLVKTCIRTFQGFHGVGVMDIADNFRIRISGSHDSIKWDLQSG- 582  
 QY 504 RAPACTLIRLVHSGVFRQDEPFIYSSHDDTI 541  
 DB 583 -----KCHHTF---NGR--RIQRETOHTQOTGSLDKV 609

AC 001277;  
 DT 15-JUL-1998 (rel. 36, Created)  
 DT 15-JUL-1998 (rel. 36, Last sequence update)  
 DT 16-OCT-2001 (rel. 40, Last annotation update)  
 DE Sulfur controller-2 (SCON2).  
 GN SCON-2.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=74-OR23-1A;  
 RX MEDLINE=95241499; PubMed=7724564;  
 RA Kumar A., Paletta J V.;  
 RT "The sulfur controller-2 negative regulatory gene of Neurospora  
 crassa encodes a protein with beta-transducin repeats.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3343-3347(1995).  
 CC - FUNCTION: NEGATIVELY REGULATES SULFUR STRUCTURAL GENE EXPRESSION.  
 CC - INDUCTION: EXPRESSED ONLY UNDER LOW-SULFUR CONDITIONS.  
 CC - SIMILARITY: CONTAINS 8 WD\_REPEATS (TRP-ASP DOMAINS).  
 CC - SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC - SIMILARITY: BELONGS TO THE MET30/SCONE/SCON-2 FAMILY OF WD-REPEAT  
 PROTEINS.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U17251; AAA68968.1;  
 DR InterPro: IPR001810; F-box.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 7.  
 DR Pfam: PF00646; F-box; 1.  
 DR PRINTS: PR00320; GPOTEINRPT.  
 DR ProDom: PD000018; WD40; 4.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 6.  
 DR PROSITE: PSS0181; FBOX; 1.  
 DR PROSITE: PSS0678; WD\_REPEATS\_1; 2.  
 DR PROSITE: PSS0082; WD\_REPEATS\_2; 6.  
 DR PROSITE: PSS0294; WD\_REPEATS\_REGION; 1.  
 DR Transcription regulation: Repeat; WD repeat.  
 FT DOMAIN 124 170  
 F-BOX.  
 FT REPEAT 292 320 WD 1.  
 FT REPEAT 332 360 WD 2.  
 FT REPEAT 372 400 WD 3.  
 FT REPEAT 411 441 WD 4.  
 FT REPEAT 453 488 WD 5.  
 FT REPEAT 528 564 WD 6.  
 FT REPEAT 576 604 WD 7.  
 FT REPEAT 616 644 WD 8.  
 SQ SEQUENCE 650 AA; 72189 MW; 7473859C99F1B028 CRC64;

Query Match 17.5%; Score 531.5; DB 1; Length 650;  
 Best Local Similarity 25.0%; Pred. No. 1.7e-32;  
 Matches 144; Conservative 76; Mismatches 180; Indels 177; Gaps 11;

QY 117 VEHLISOMCHYGHINSYKPMLODFITALPARGIDHIAENILSYDAKSLCAELVC 176  
 DB 99 LOGILSQLEPQLSFVSREVENALKIDFISALPV-----ELAQKVLGYLDTVSLTKAAQVS 154  
 QY 177 KENVRYTSDGMLMKLIERVVRIDSLMRGLAERGGVLYFKNK-----P 222  
 DB 155 QKRRTIADSDAVVWVRCQEHVNRK-----CTKCGWGLPLERKKILNRYRROLAKGP 208  
 QY 223 DGNA----- 222  
 DB 209 QGAVTELADSHDSQDSVNOHGKRPAAEAERDDPIKKRCMAAASRAVYTOPKTRSWKA 268



RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA  
 CC REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE  
 CC POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD  
 CC DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONJUGATION AND  
 CC VARIOUS ASPECTS OF SPOROGATION. REQUIRED FOR HTAI-HTBI LOCUS  
 CC TRANSCRIPTION ACTIVATION (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X96763; CAA65538.1; -  
 DR InterPro: IPR001810; F-box.  
 DR Pfam: PF004400; WD40; 7.  
 DR Pfam: PF00446; F-box; 1.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 1.  
 DR SMART: SM00256; FBOX; 1.  
 DR SMART: SM00320; WD40; 7.  
 DR PROSITE: PS50181; FBOX; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 4.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 4.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW Cell division; Mitosis; Sporulation; Repeat; WD repeat.  
 FT DOMAIN 212 258 WD 5.  
 FT REPEAT 322 351 WD 1.  
 FT REPEAT 363 391 WD 2.  
 FT REPEAT 403 431 WD 3.  
 FT REPEAT 442 468 WD 4 (POTENTIAL).  
 FT REPEAT 478 506 WD 5.  
 FT REPEAT 519 549 WD 6.  
 FT REPEAT 561 589 WD 7.  
 SO SEQUENCE 684 AA; 76090 MW; 3DD65DB31293B107 CRC64;  
 Query Match 15.7%; Score 477.5; DB 1; Length 684;  
 Best Local Similarity 25.4%; Pred. No. 2.1e-28;  
 Matches 136; Conservative 95; Mismatches 184; Indels 121; Gaps 16;

QY 446 GSSDNTIRIMDIEGACGLVLEGHELYNCIRFDNRKRYSGAYDGKIRKVDLVALDPRA 505  
 DB 579 AADATLRIMDAKTEFLRSKLGHCALTCFEGHDLRVVSGS-EKMLKLMV----- 629  
 QY 506 PAGTLCRLTVSHSGRVFLQPD-----EQIYSSHDITIL-IMPFLNDPAQ 553  
 DB 630 EKGTRARDLLSPVTGIGWQRPDYRCVAAVGRILNEDEGETFELIDF-SQPLIQ 684  
 RESULT 12  
 ID HET1\_PODAN STANDARD; PRT; 1356 AA.  
 AC Q00808;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Vegetable Incompatibility protein HET-E-1.  
 GN HET-E1.  
 OS Podospora anserina.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariales; Lasiosphaeriaceae; Podospora.  
 OX NCBI\_TaxID=5145;  
 RN [1]  
 RP MEDLINE=96009891; PubMed=7557402;  
 RA Saude S., Turc B., Begueret J.;  
 RT "A gene responsible for vegetative incompatibility in the fungus  
 RT Podospora anserina encodes a protein with a GTP-binding motif and G  
 RT beta homologous domain."  
 RL Gene 162:135-139(1995).  
 CC -1- FUNCTION: RESPONSIBLE FOR VEGETATIVE INCOMPATIBILITY THROUGH  
 CC SPECIFIC INTERACTIONS WITH DIFFERENT ALLELES OF THE UNLINKED GENE,  
 CC HET-C.  
 CC -1- SIMILARITY: CONTAINS 10 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: CONTAINS 1 NACHT DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L28125; AAA85775.1; -  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF004400; WD40; 10.  
 DR PRINTS: PR00320; GPROTEINBRPT.  
 DR PRODOM: PD000018; WD40; 10.  
 DR SMART: SM00320; WD40; 10.  
 DR PROSITE: PS50837; NACHT; 1.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; 10.  
 DR PROSITE: PS50082; WD\_REPEATS\_2; 10.  
 DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
 KW GTP-binding; Repeat; WD repeat.  
 FT DOMAIN 294 629 NACHT.  
 FT NP\_BIND 300 307 GTP (POTENTIAL).  
 FT REPEAT 839 869 WD 1.  
 FT REPEAT 881 911 WD 2.  
 FT REPEAT 923 953 WD 3.  
 FT REPEAT 965 995 WD 4.  
 FT REPEAT 1007 1037 WD 5.  
 FT REPEAT 1049 1079 WD 6.  
 FT REPEAT 1091 1121 WD 7.  
 FT REPEAT 1133 1163 WD 8.  
 FT REPEAT 1175 1205 WD 9.  
 FT REPEAT 1217 1247 WD 10.  
 SO SEQUENCE 1356 AA; 149765 MW; 965FB319844E0651 CRC64;  
 Query Match 15.0%; Score 455.5; DB 1; Length 1356;  
 Best Local Similarity 26.0%; Pred. No. 2.3e-26;  
 Matches 164; Conservative 102; Mismatches 220; Indels 145; Gaps 30;

```

OY 28 EPPRKIIPEKNSLRQYINSCAR-ICLNQETVCLASTAMK-----TENCYAKTKLANGTSS 81
DB 581 EMDVDSDDESELEIYKLCGSLIIRERIVYVHOSAKDFLLGTMSDKANSKASOEARE 640
OY 82 MIVPR-----ORKISASYEKEKE--LCYKFEBSSESQOFEVHL--ISQMCHY 127
DB 641 LVEPTGIEDVSYLIFWRSILNVMSSQKLRDRDYCLINAGFLIDNRVDPDPPLATVRYSCTI 700
OY 128 OHGHINYLKP-----MLQRO-----FT-----ALPANGLDHI--AEN 159
DB 701 WIDLRLDVLSTSSKWHVLLQDDGDHRLFTYKLYLWLEALSLRLAP--EGVAINROLSS 759
OY 160 ILAYLDAKSCIAELVCKEYRVATSDGMLMKLIER---NVRTDSGLRDLAERRGMOYVL 216
DB 760 LLGHTINGRLIA---LYRDCYRF---ALSTRMIEKAPLQATYSAL-----V 800
OY 217 FKNKPPDGNAPNPSFYALY---PKLIIDLEITESWRCGRHSL----- 257
DB 801 F-----APTDSMIKKIKKEPQWISITISVEAEEMNACTOTLEGHSSVLSVAFSAD 852
OY 258 -QRIHCS-----ETSKG-----VYCLQY--DDQKIVSGLRDNTIKIMDK 294
DB 853 GQRYASGSDDKTIKIMDTASCTGTOTLEGHGSYSAFSPDRNRVAGSGDDTIKIMDA 912
OY 295 NTECKRRIILGHTGVSCLQY--DERVLIIGSSDSITVAVDVTGEMLNTLIIHCEAVLH 352
DB 913 ASGCTGTOTLEGHGRVGSVAFSPGQGRVAGSGDDTIKIMDAAGCTGTOTLEGHGSVLS 972
OY 353 LREN-NGMMWTC-SKDRSIAVMDASPTDITLRVLYVGHRAVNVVP--DDKIYVAGS 408
DB 973 VAFSPDQGRVAGSGDDTIKIMDTASCTGTOTLEGHGSVSAFSPGQGRVAGSGD 1029
OY 409 DRTIKVWNTSTCEVPTFLNGHKGRIACLOXR--DRLVAGSGSDITFLMDIEGCACRLV 466
DB 1030 DKTIKIMDTASCTGTOTLEGHGWGVSVSPDQGRVAGSGDDTIKIMDAVSTGTOTL 1089
OY 467 EGHEELVRCIRF--DNKRIYSGAYDGKIKWMDLVAALDPRAPACTICLRIIVHSGVRF 524
DB 1090 EGHGDSVMSVAFSPDQGRVAGSGDDTIKIMD-----AASGT-CTGTLEGHGWVHS 1140
OY 525 LQF--DEQIVSSSHDITILIMDLNDPMAQ 553
DB 1141 VAFSPDQGRVAGSGDDTIKIMDAAGCTGTQ 1171

RESULT 13
POP1_SCHPO STANDARD; PRT; 775 AA.
AC P87060;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE WD-repeat protein popl.
GN POP1 OR SPBC262.18.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OC NCHI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RX MEDLINE-97347242; PubMed-9203581;
RA Komitani K., Toda T.;
RT "Fission yeast WD-repeat protein popl regulates genome ploidy through
RL ubiquitin proteasome-mediated degradation of the CDK inhibitor Rum1
and the S-phase initiator Cdc18."
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RX MEDLINE-21848401; PubMed-11859360;

```

```

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Mblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Koben J., Grymopre B.,
RA Welljens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Mambuti R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaune V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerniti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RA "The genome sequence of Schizosaccharomyces pombe."
RT Nature 415:871-880(2002).
RN [3]
RP SUBUNIT.
RC STRAIN-972;
RX MEDLINE-99144318; PubMed-9990507;
RA Komitani K.-I., Ochocotore I., Toda T.;
RT "Two F-box/WD-repeat proteins Pop1 and Pop2 form hetero- and homo-
RT complexes together with cullin-1 in fission yeast SCF (Skp1-cullin-1-
RT F-box) ubiquitin ligase."
RL Gene Cells 3:721-735(1998).
CC - FUNCTION: Involved in maintenance of ploidy through proteasome
CC dependent degradation of CDK inhibitor rum1 and S-phase initiator
CC cdc18. Functions as a recognition factor for rum1 and cdc18, which
CC are subsequently ubiquitinated and targeted to the 26S proteasome
CC for degradation.
CC - SUBUNIT: Homodimer and heterodimer with pop2. Binds to pcu1 and
CC cdc18.
CC - SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC - SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: Y08391; CA65671.1; -
DR EMBL: AL022103; CA17898.1; -
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 7.
DR Pfam: PF00646; F-box; 1.
DR PRINTS: PR00320; GPROTEINRPT.
DR ProDom: PD000018; WD40; 3.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 6.
DR PROSITE: PS00181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 3.
DR PROSITE: PS50082; WD_REPEATS_2; 4.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
FT DOMAIN 298 345 F-BOX.
FT REPEAT 444 472 WD 1.
FT REPEAT 484 538 WD 2.
FT REPEAT 575 603 WD 3.
FT REPEAT 615 645 WD 4.

```

```

RC SEQUENCE FROM N.A.
RP STRAIN=S288c / AB972:
RX MEDLINE=95400292; PubMed=7670463;
RA Murkaml Y., Naitou M., Hagihara H., Shibata T., Ozawa M.,
RA Sasunuma S.-I., Sasunuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae."
RT Nat. Genet. 10:261-268(1995).
RN [3]
RP SEQUENCE OF 1-579 FROM N.A.
RC STRAIN=S288c / AB972:
RA Barrell B.G., Churcher C., Raeburn M.A.;
RL Submitted (SEP-1994) to the EMBL/Genbank/DBJP databases.
CC -! FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA
CC REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE
CC POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD
CC DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONJUGATION AND
CC VARIOUS ASPECTS OF SPOULATION. REQUIRED FOR HTAI-HTBI LOCUS
CC TRANSCRIPTION ACTIVATION.
CC -! SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -! SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X05625; CAA29113.1; -
DR EMBL; D50617; BAA09229.1; -
DR EMBL; Z46255; CAA86341.1; -
DR PIR; A26867; A26867.
DR SGD; S0001885; CDC4.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001880; WD40.
DR Pfam; PF00400; WD40; 7.
DR Pfam; PF00646; F-box; 1.
DR PRINTS; PR00320; GPROTEINBPT.
DR ProDom; PD000018; WD40; 4.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 6.
DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 4.
DR PROSITE; PS50082; WD_REPEATS_2; 5.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Cell division; Mitosis; Sporulation; Repeat; WD repeat.
FT DOMAIN 272 319 F-BOX.
FT REPEAT 380 408 WD 1.
FT REPEAT 420 449 WD 2.
FT REPEAT 461 493 WD 3.
FT REPEAT 528 556 WD 4.
FT REPEAT 568 598 WD 5.
FT REPEAT 630 658 WD 6.
FT REPEAT 669 698 WD 7.
FT CONFLICT 460 460 K -> E (IN REF. 1).
SO SEQUENCE 779 AA; 86089 MW; 0348P2F8FA78F3BC CRC64;

Query Match 13.2%; Score 399; DB 1; Length 779;
Best Local Similarity 24.3%; Pred. No. 2e-22;
Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

QY 56 TVCLASTAMKTENCVAKTAKLANGTSSMIVPKORLASYEKER-----ELCYKFEQW 108
DB 194 TTPLAKTKTKTINN-----NNNTADIIESKDSITISEYLSDEFSAINNLTPLAYFK-- 244

QY 109 SESDQVEFEVHLISOMCHYOHGHSNVLKPLLODFDTLTPARGLDHTAENIISYLDAKS 168
DB 245 -----NLIRLVANMRSELSDIGTILKDKLAKRDLITSLPF-----RISLKITFYLOPFD 294

QY 169 LCAAEIVCKEWR-VTSDGMLMKLLI--ERRVVRDTSIMRGIAERGGQVLFKKKPPDGN 225

```



Matches 146; Conservative 92; Mismatches 226; Indels 163; Gaps 25;

```

OY 14 KFMNSSEED--CNNGPEPRKI-----IPEKNSLRJOTYNSCARCL-----NOETVCL 59
Db 96 KFMSEIENVATKCLDHPSPDPDPDFSIUVPQNFL-YSHSSLPKXISIDRNNRKLD 154
OY 60 ASTAMKTENCVAKTUKLANGISSMTVPKOKIASYER-EKELCAYFPEWSSDQVEYE 118
Db 155 NSISSENSNPPSPRPV--DISNTYSPSPKPSISELDELNIQSIQVPEFEDJPEGIO-SYAF 211
OY 119 HLIQSMCHYQGH--INSYLRMLORDFTALPARGDLHRENIILSYIDAKSLCAELVCK 177
Db 212 FQLRSCROGMRILLNCEPELKKDILSNPF--SIYOSTILNNDHSPLSCRYSVP 267
OY 178 EMYRV-----TSDGMLMKULIRMYRTTSLNKGLEERGGWGOYLFPKKNPUGGNAP-----N 229
Db 268 TWNRLLDVHTS--YWKHMF-----SLFEGQINENDW--KYANPNLNPPELLHND 312
OY 230 SFYRALYKRIIODIETIESNMRCGRHSIOR-----259
Db 313 QISDDYPEPEIK-----RHFLNRKRMLFPSPISPHLSFPIHVNFMTSLLLH 360
OY 260 -----HCRBETSK-GYCCQYODOKIYVSLRONTIKIMDKNT 236
Db 361 KDRITTTSGSGTQIHNATGYLVARLEGHKEGEMAWAKIHENTVSSIDKTYRWAMTKE 420
OY 297 LECKRIILGHMGSVLCLQY-----DERVIITGSSDSTVAVWVONTGEML 340
Db 421 AKCTIEFEGHSIIIRCTELIIPVRLIRHGEIVEPDPQPIYVSGSRDHTLRW-----472
OY 341 NTLIHCEAVLHLRPNNSMNVGSKDRSIAVMDASPTDLIRVYAGHGAANYVDEDD 400
Db 473 -----KLKRNDDPYLPDNTNISIDRNE-KNYPEV---HTLIGHDSVRTISGYG 517
OY 401 KYIVASAGDKTIYKWNSTJCFEVRTLNGHKRGIACLQYRD--RIIVSGSSDNTIRLMDIE 458
Db 518 DILVSGSDSSIRIKRWSTJGCLTHLGHSHLRITSYLVEPERNICISGSMKDSIRVWDL 577
OY 459 CGACILRVGHEEHLVRCIRFPNKRIRVSGAYVQIKVMDVLAALDPRAPACTILRLTYEH 518
Db 578 TGTCTYVLEGHDAFVTLLNVPQNMLJSGSADSTIRIMD-----LNTGKP-----LMTLPSN 628
OY 519 SGRVRLQDFDFQIYSSSHDTITLIMD 545
Db 629 SGYISSFVSDERKII-SGNDSSVYKLM 654

```

Search completed: February 20, 2003, 09:54:19  
Job time : 19 secs



ADDRESSEE: David A. Jackson, Esq.  
 STREET: 411 Hackensack Ave, Continental Plaza, 4th  
 STREET: Floor  
 CITY: Hackensack  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07601

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/914,999  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Jackson Esq., David A.  
 REGISTRATION NUMBER: 26,742  
 REFERENCE/DOCKET NUMBER: 601-1-078  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 201-487-5800  
 TELEFAX: 201-343-1684  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2237 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 HYPOTHEICAL: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Dictyostelium discoideum  
 US-08-914-999-7

Query Match 3.8%; Score 82.2; DB 4; Length 2237;  
 Best Local Similarity 47.4%; Pred. No. 3e-16;  
 Matches 286; Conservative 0; Mismatches 308; Indels 9; Gaps 1;

QY 959 TGCATGCAAGGATTCCTACAGAGCCATACAGGTCCTGCTCCAGTATGATG 1018  
 DB 1430 TGGATGTCGTCACAACTTGAAGGTCTATGAAGGTCCAGTTGATTAATG 1549  
 QY 1019 AGAGATGATCATACAGATCATGATTCACGGTCAGAGTGGATTAATACAG 1078  
 DB 1550 ATCAATATTTGTTAGTGTTCATGATCAATTAATTAAGTTGGATTAAGAAT 1609  
 QY 1079 GTGAAATGCTAAACAGCTGATTCACCATTTGAAAGCATCTGCTGCTTCAATA 1138  
 DB 1610 TAAATGATTTTACCTTGAAGGTCTATGAATTAACCTGTCATACGATTCATTAATG 1669  
 QY 1139 ATGCGATGATGTCACCTGTCACAAAGATCCCTGATGATGAGATATGAGCTCC 1198  
 DB 1670 ATAAATATTTGTTAGTGTTCCTGTCGACAAAGTCAAAAGTTGGATTTGAAACTT 1729  
 QY 1199 CACTGACATTAACCTCCGGAGGTCGTGTCGACACCGAGCTGCTGCTCAATTTGTAG 1258  
 DB 1730 TGGAAATTAATATACCTTGAAGGTCAATGACCA-----GCCGTCAAAACACTTT 1780  
 QY 1259 ACTTGATGACAGATACATTTCTGTCATCTGGGATAGAACTATTAAGATATGAAACA 1318  
 DB 1781 GTATATCTGTCGACATTTATTTAGTGTTCATTAATGATTAACATATCAAGTTGGGAT 1840  
 QY 1319 CAAGTACTGTGAATTTGTAAGACCTTAAATGACACAAAGAGGATGCTGTTGC 1378  
 DB 1841 TGAATACTTTGTTGTAATACACTCTAAAGGTCACTAATAATGAGTCAACCACTATCT 1900  
 QY 1379 AGTACAGGACAGGCTGATGAGTGGCTATGACAAACATATGAGATATGAGGACA 1438  
 DB 1901 GTATATTTAGTACCAATCTTACAGTGGCTCTATGATTAATAAATATAGAGTTTGAATT 1960  
 QY 1439 TAGAATGTGTGATGTTTACGAGTGTAGAGGCAATGAGGAATTTGCTGTTGATTC 1498  
 DB 1961 TAAAGATTTAGATGTTCCGCTACTTTAAGAGCCCATGATGATGTTGAACATATGG 2020  
 QY 1499 GATTTGATTAACAGAGATATGAGTGGGCTATGATGAAATTAAGTGGGATC 1558  
 DB 2021 TAAATTTGATTAATATATTTACTGCTAGTACGATTAATACATTTAAATTTGGATT 2080  
 QY 1559 TTG 1561  
 DB 2081 TAG 2083

RESULT 3  
 US-08-899-578-1  
 ; Sequence 1, Application US/08899578  
 ; Patent No. 6087153  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Greenwald, Iva  
 ; APPLICANT: Hubbard, E. Jane  
 ; TITLE OF INVENTION: SEL-10 AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Cooper & Dunham LLP  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/899,578  
 ; FILING DATE: 24-JUL-1997  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: White, John P.  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 0575/53200/JPW/AKC  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 278-0400  
 ; TELEFAX: (212) 278-0525  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2481 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 91..1854  
 ; US-08-899-578-1

Query Match 3.1%; Score 67.2; DB 3; Length 2481;  
 Best Local Similarity 50.6%; Pred. No. 2.6e-11;  
 Matches 162; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

QY 1273 TATATGTTTGCATCTGGGATTAAGTATTAAGATGAGACAACTGTTGTA 1332  
 DB 1018 TATATGTTTACGGGGCTCACTGATTAAGTGAATTTGAGTATGATGATGTTCA 1077  
 QY 1333 TTTTGAAGACCTTAAATGACACAAAGAGCAATTCCTGTTGCTACAGAGGACAG 1392  
 DB 1078 CTCTCTTCATACACTTAAAGACATCTTCACTGTTGATGATGATGCTATGCTGTTCC 1137  
 QY 1393 CTGTGTGATGAGTCTCATCTGACAACTATACATTAATGAGACATGATGTTGCA 1452  
 DB 1138 ATACTGTTGATGAGACAGATACCACTCTTCTGATGAGGACGTAGAAATCCGAGCT 1197  
 QY 1453 TGTTCAGATGTTTGAAGGCAATGAGGAATTTGTCGTTGATTCATTAACAG 1512



US-08-232463-14/c  
 RESULT 6  
 ; Sequence 14, Application US/08232463/c  
 ; Patent No. 5670367  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DORNER, F.  
 ; APPLICANT: SCHEFFELINGER, F.  
 ; APPLICANT: FALKNER, F. G.

[illegible]



Db 1112 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1171

US-08-664-596B-1

Query Match	1.8%;	Score 38.6;	DB 1;	Length 433;
Best Local Similarity	51.4%;	Pred. No. 0.019;		
Matches	89;	Conservative	0;	Mismatches 84;
			Indels	0;
			Gaps	0;





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 09:55:06 ; Search time 22 seconds

(without alignments)  
2486.386 Million cell updates/sec

Title: US-09-601-168b-2

Perfect score: 3034

Sequence: 1 MDPAEAVLQERALKRFMNSSE.....PAAQAEPPSPSRITYTISR 569

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

PIR-73:\*  
1: pirl:\*  
2: pirl:\*  
3: pirl:\*  
4: pirl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2597	85.6	518	2	B48088
2	1635.5	53.9	701	2	T16607
3	690	22.7	506	2	T50211
4	590.5	19.5	605	2	T38932
5	545	18.0	640	2	S49932
6	531.5	17.5	650	2	T46660
7	520	17.1	579	2	T22703
8	519.5	17.1	267	2	S62507
9	455.5	15.0	1356	2	T18521
10	453	14.9	775	2	T45136
11	413.5	13.6	1227	2	AE1810
12	399	13.2	779	2	S56245
13	396.5	13.1	703	2	T43557
14	378.5	12.5	1189	2	A12493
15	375	12.4	1747	2	AC1842
16	374	12.3	1526	2	AC2239
17	373.5	12.3	1258	2	A12155
18	373.5	12.3	1683	2	AE2071
19	361.5	11.9	677	2	AE1861
20	358.5	11.8	559	2	AB2202
21	354	11.7	409	2	S61133
22	354	11.7	410	2	S48052
23	341	11.2	515	2	S19487
24	339.5	11.2	777	2	T41075
25	337	11.1	1146	2	A55532
26	336	11.1	676	2	AH2195
27	334	11.0	589	2	AG2400
28	333.5	11.0	1711	2	AD1842
29	332.5	11.0	317	2	T46032

30	330	10.9	333	2	G85034	probable WD-repeat
31	320	10.5	934	2	AG1889	WD-40 repeat prote
32	318	10.5	422	2	A56640	CDC4 repeat unit-c
33	313.5	10.3	376	2	T19266	hypothetical prote
34	312.5	10.3	714	2	S56893	hypothetical prote
35	312	10.3	786	2	AC2375	WD-40 repeat-prote
36	312	10.3	876	2	T51507	WD-40 repeat protei
37	307.5	10.1	1693	2	S76086	beta transducin-11
38	306.5	10.1	598	2	AE2415	WD-repeat protein
39	306	10.1	1189	2	AH2154	WD-repeat protein
40	305	10.1	704	2	S33263	transcription init
41	299.5	9.9	1194	2	T03818	apoptotic proteina
42	298	9.8	265	2	AF1890	WD-repeat protein
43	297.5	9.8	357	2	A12099	WD-40 repeat prote
44	297	9.8	502	2	T41148	trip-asp repeat con
45	296.5	9.8	304	2	AG1837	WD-40 repeat prote

## ALIGNMENTS

### RESULT 1

B48088 beta-transducin repeat-containing protein - African clawed frog

N:Alternate names: beta-trcp

C:Species: Xenopus laevis (African clawed frog)

C>Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 21-Jul-2000

C:Accession: B48088

R:Spevak, W.; Keiper, B.D.; Stratowa, C.; Castanon, M.J.

Mol. Cell. Biol. 13, 4953-4966, 1993

A:Title: Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase

A:Reference number: A48088; MUID:93530289; PMID:8393141

A:Accession: B48088

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-518 <SPe>

A:Cross-references: GB:M98268; NID:9295542; PUID:AAA02810.1; PID:9295543

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

C:Keywords: duplication

F:431-462/Domain: WD repeat homology <MD1>

Query Match	Score	Length	DB ID	Description
Best local similarity	91.6%	Pred. No. 1.2e-193		
Matches 468; Conservative 7; Mismatches 8; Indels 30; Gaps 1;				
18	SSEREDCNNGEPPRKIIPEKNSLRQTYNSCARCLNQETVCLASTAKTECVAKTKLAN	77		
Db	13 ASEREDCNNGEPPRKIIPEKNSLRQTYNSCARCLNQETVCLASTAKTECVAKTKLAN	42		
QY	78 GTSMIVPKRKISASYEKEKELCVYFFQMSRSDQVEVEHLISOMCHYOHGHTSYLK	137		
Db	43 GTSMIVPKRKISASYEKEKELCVYFFQMSRSDQVEVEHLISOMCHYOHGHTSYLK	102		
QY	138 PMLORDTITLALPARGDHIENITLSTYDAKSLCAELVCKEMRYVSDGLMKLTERVY	197		
Db	103 PMLORDTITLALPARGDHIENITLSTYDAKSLCAELVCKEMRYVSDGLMKLTERVY	162		
QY	198 RTDSLWGLAERKRGWGYLFKNRPDGNAPNSFYALPKIITQDIETTESNRCGRHSI	257		
Db	163 RTDSLWGLAERKRGWGYLFKNRPDGNAPNSFYALPKIITQDIETTESNRCGRHSI	222		
QY	258 QRIHCRSETSKGYVCIQYDQOKIVSGLRDNTIKIMDKNTLECKRILGHGHSVLCQYDE	317		
Db	223 QRIHCRSETSKGYVCIQYDQOKIVSGLRDNTIKIMDKNTLECKRILGHGHSVLCQYDE	282		
QY	318 RVITIGSSDSSTVRYWVDNTEGMLNTLIHCEAVLHLRFNNGMAYTSCDKRSIAVWMAAP	377		
Db	283 RVITIGSSDSSTVRYWVDNTEGMLNTLIHCEAVLHLRFNNGMAYTSCDKRSIAVWMAAP	342		
QY	378 TDTILRLVILGHRAAVNVVDEDDKIIVSASGDRITIKWNTSTGEFVTLNGHKGRLACIQ	437		
Db	343 TDTILRLVILGHRAAVNVVDEDDKIIVSASGDRITIKWNTSTGEFVTLNGHKGRLACIQ	402		

QY 438 YRDRLVYSSGSSDNTLRIMDLICGACGLVYLGSHHELVYLCIFPNDKRIVYSGAYDCKITVMDL 497

Db 403 YRDRLVYSSGSSDNTLRIMDLICGACGLVYLGSHHELVYLCIFPNDKRIVYSGAYDCKITVMDL 462

QY 498 VVALDPRAPACTLCRLTLVHSGVFLQDPERQYVSSSHDTITLIMDFLNDP 550

Db 463 VVALDPRAPACTLCRLTLVHSGVFLQDPERQYVSSSHDTITLIMDFLNDP 515

RESULT 2  
M16607

hypothetical protein K10B2.1 - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C/Accession: T16607  
R/Miller, N.  
submitted to the EMBL Data Library, June 1995  
A/Description: The sequence of C. elegans cosmid K10B2.  
A/Reference number: Z18545  
A/Accession: T16607  
A/Status: preliminary; translated from GE/EMBL/DBD  
A/Molecule type: DNA  
A/Residues: 1-701 <MIL>  
A/Cross-references: EMBL:U28730; NID:9860694; PID:9860695; PIDN:AAAG8258.1; CESP:K10B2.1  
A/Experimental source: strain Bristol N2  
C/Genetics:  
A/Gene: CESP:K10B2.1  
A/Introns: 78/3; 125/1; 183/2; 281/3; 404/3; 551/3; 668/3

Query Match	53.9%;	Score 1635.5;	DB 2;	Length 701;
Best Local Similarity	57.7%;	Pred. No. 8.1e-119;		
Matches 328;	Conservative 69;	Mismatches 116;	Indels 55;	Gaps 8

QY	31	RIIIEKNSIRQTVNSCARLILNOETVCLASTAKTECVAKTYLAN-----GTS	80
Db	2	RREBRGKALRQGRHARDGCSIALQVCSV---IERCF--TAVSNPIFFLESTFSPVF	55
QY	81	SMIVQROR-----KHASYEKEKELCYVFPQWSESDQVEVEHLISOMOHYGHNSY	135
Db	56	SPLFPSRNTQIFLAYSRSFSSFEVL-----KSEHEDQDPMQIVIRLSHYOLGAVDNF	110
QY	136	LKPMLOPFTALPARGLDHTIAENILSYLAKSLCAALCYCKEYRVYSDGMMLKILIER	195
Db	111	IRPMLOPFIENLPA---HLVEILFENVNSDSIKSCEEVSTSWRCALARGQHMKLIER	166
QY	196	MVRTSLMRGLAERGMGOYL-----FNKPPDGAAPNSFTYRAL	235
Db	167	NVRSDSLMWGLSEKROMDKPLNISRDMSVRIICEKNYDVNIRKDLQDILMHVEYSKI	226
QY	236	YPKIIODIETIESMWRGCRHSIORHCSEFTSKGVYCYLOVDQDKIVSGILNDNTKILMDKN	295
Db	227	YPKIIRIDHINIDNMKKGNVEMTRIINOQSEMSKVYCYLOVDQDKIYSGILNDNTKILMDK	286
QY	296	TLECKRILTGHTGSVLCLOYDERVIITGSSDSYRVWVDNTGEMLNTLILHCEAVLHLRF	355
Db	287	DYSCSRIISGHTGSVLCLOYDNRIYISGSSDAPYRVWVDYETGECIKTILIHCEAVLHLRF	346
QY	356	NGMAVATCSKRSRJAIVMDMASPTTITRLRYLGVHRAVNVYDEDDKYIYASASDRITIKW	415
Db	347	ANGIMATVCSKDRSIAVVDMSVPRDITIRKVLGVHRAVNVYDEDDRYIYASASDRITIKW	406
QY	416	NTSTCFEVRTLNGHKRGIACLQYRDRLVYSGSSDNTIRLMDIECGACILRYLKEHEELVRC	475
Db	407	SMDTLEFVRTLAGHRGRIACLOYRGRILVYSGSSDNTIRLMDIHSGYCLAVLBHEELVRC	466
QY	476	IRPNKRTKVSATYDGKIKVMDLYAALDPRAPACTICLTRLVEHSGVFRLODEFOIYSS	535
Db	467	IRPEKRTVSGAYDGKIKVMDLOAALDPRALSEIICISIVGHTGAVFRLODDFOIYSS	526
QY	536	SHDITILIMPLNDPAQAEPSPSPRT	563
Db	527	SHDITILIMPLNDP-----PSGIPSPST	549

RESULT 3  
T50211

### RESULT 3

WD-repeat protein [imported] - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 02-Sep-2000  
 C:Accession: T50211  
 R:McGouall, R.C.; Rajadream, M.A.; Barrell, B.G.; Brown, S.; Murphy, L.; Jones, L.,  
 submitted to the EMBL Data Library, January 2000  
 A:Reference number: Z25046  
 A:Accession: T50211  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-506 <MCD>  
 A:Cross-references: EMBL:AL136538; PIDN:CAB6464.1; GSPDB:GN00066; SPDB:SPAC30.05  
 A:Experimental source: strain 972h(-); cosmid c30  
 C:Genetics:  
 A:Gene: SPAC29E6.01; SPDB:SPAC30.05  
 A:Map position: 1  
 A:Introns: 43/1; 74/3  
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match	22.7%;	Score 630;	DB 2;	Length 506;
Best Local Similarity	30.4%;	Pred. No. 1.3e-45;		
Matches 163;	Conservative 104;	Mismatches 167;	Indels 102;	Gaps 15;

QY	67	ENCVAK----	TKLANGTSSMVPQRFKLTASBYEKELCVKFFEDMSSDOVEVEHLIS	1222
Db		1	1	1
QY	123	QMCYHOGHINSYKPMLOQDFITALPARGLOHIAENILSYLDANKSLCMAELVCKEWRV	1822	
Db		1	1	1
QY	183	TSDGMIKMKLI-----	EBWRTDSIMWG	213
Db		1	1	1
QY	107	LDDPEIKMALLYMKGWVYNNENVLNEEPAMRRHKFPQPFENFLKQOONITIGPYGTMFLPQ	166	
Db		1	1	1
QY	214	QYLFKNKPPGDNAPNSFYALYPKIIIDLETESNMRCGRHSLOHICRS-----	264	
Db		1	1	1
QY	167	QEIF-----	DNGRPLINWSTLY-----KEHAHLDSSWHRGRLVSTFNNPISREPADQF	217
Db		1	1	1
QY	265	-FTSKGVYCLQYDOCKVYSGLRDNTKIKMDKNMLDECKRLIGHGSCSYCLQYDGR--VII	321	
Db		1	1	1
QY	218	RATLDVYCYQYDEIMVYSSKDRITYSMDVNSKRTLYKTLGHSSSYCLDPCRRNLLY--	277	
Db		1	1	1
QY	322	TGSSDSYIRVADVNTGEMTLTLIHCEAVLHLRFNNGMAYTCKSDRSIAYW--DMASPTD	379	
Db		1	1	1
QY	278	SGSSDSITIIDMOMNRRLPYVYCGHDVNLGVVSENYLSSSDHDPARVWRDATSPAE	337	
Db		1	1	1
QY	380	ITLRRLVILGHAAVNVYDFDK--VTYSASGVRTIKVNTSTCEFTVINGHKGATCQ	437	
Db		1	1	1
QY	338	ACM--HYLRGHLASVNSVOYSKKTGLTVTASSDRTLRWDITTGHCIRIHHOQGIACQ	396	
Db		1	1	1
QY	438	YRDLRYVSGSSDNTIRLMDIEGCACLRVLEGHEELVRCIRPDNRKIYASGADKIRYWDL	497	
Db		1	1	1
QY	397	YNGKFIVYSGSDLLTIRIFEFSAGKILRMLOGHEDLITFVRFNDEKIVYSGGYDGIYRW--	455	
Db		1	1	1
QY	498	VAAIDPAPAGTICTLRIVH-----	SGVRERLOCFEPIYSSSDDTIILWDF	546
Db		1	1	1
QY	456	-----	FNTGDGCHVLAHNSRNSRYFGIOLDFPHRRITIACTHSSSETLWNP	497
Db		1	1	1

## RESULT 4

proteasome shuttle metalloprotein - fission yeast (*Schizosaccharomyces pombe*)  
C.Species: Schizosaccharomyces pombe  
C.Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 26-May-2000  
C.Accession: T38932  
A.Submitter: K.J. Church, C.M.; Wood, V.; Bartell, B.G.; Rajandream, M.A.  
A.Submitted to the EMBL Data Library, April 1997  
A.Reference number: Z21818  
A.Accession: T38932

A>Status: preliminary; translated from GR/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-605 <BAD>  
 A:Cross-references: EMBL:294864; PIDN:CA08168.1; GSPDB:GN00066; SPDB:SPAC57A10.05C  
 A:Experimental source: strain 972h; cosmid c57A10  
 C:Genetics:  
 A:Gene: SPDB:SPAC57A10.05C  
 A:Map position: 1  
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 19.5%: Score 590.5; DB 2: Length 605;  
 Best Local Similarity 29.0%: Pred. No. 8.7e-38;  
 Matches 156; Conservative 86; Mismatches 189; Indels 107; Gaps 12;

```

QY 80 SSMIVYKORRLKLSAYEKEKELCVYFEOWSE---SDQVEVEHLISOMCHYGHINSYL 136
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 46 SSM---HNELSSEKSRQREVAWVAAPSEASCSERKIALQILNCGSSLSLFFASSTL 101
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 137 KPMLODFTTALPARGLDIAENILSYLDKSLCAAEVCKEYRTSDGMLKILIERM 196
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 102 DSLVRLDFSLPLPV---ELSPRLSFLDARSICQAQVSKHMKELADDVIMHRCCEH 157
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 197 VRTDSLIRGLAERRGWOYLFRN----- 219
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 158 INRK-----CEKCGWGLPULLENKNTLYAAKASTQKREYELTRKVDQAHESPVKKAKLD 211
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 220 -----KPPDGNAPPNSF-----YRALYKIIIDTETTESNMRGRLSLQ 258
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 212 DVTSSNEETISSVCPSPNSDSKFLPPTKTRPKVEYAEKRC---VECNMRHGR----- 263
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 259 RIHCR-----SETSKGYCCIQYDQKIVSGLRDNTIKIMKNTLECKRIILTGHTGSLCQ 314
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 264 ---CROVVLSGHSDGWCQQLVKNILASGSDATIRLMIAFQOYVALLLEGHSSEVTCQ 320
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 315 YDERVITIGSSDSTVRYVMDVNTGEMLNTLIHCEAVLHLRFNNGMATVCSKRSIAVMD 374
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 321 PDCKLISGSMKTRIKIMRTSECSILHGHDSVLCILFTSTLVSSADCTVAKLHMF 380
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 375 ASPDTITLRLVYGHRAAVNVDF--DDKYIVASAGRTIKVWNTSTCEFTVTLNGHKG 432
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 381 SGGRKRIITLR---GHTGPNVSRIIDRGVLVSGSDSTIKIWSLETNCLMFTFSIHDP 436
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 433 IACLOYRDLRVYVSGSDNTIRLMDIFCGACLRVLEGEHELVCIRPDNKRIYSGAYDGI 492
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 437 VOSLALADNRLFCSDIGDTIKQMDIEKKKCVHTLFEGHIGVEIMADHRLISGADGVV 496
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 493 KVMVDVAALDPRAPACTLCLRLTVESHGVRFLQDEFOIVSSSHDITLLIMDELNDP 550
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 497 KWEACE-----CVHTLKNHSEFVTSVALGDCEVSGSEDKILYLMFNAP 543
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

# RESULT 5

S49932  
 MET30 protein - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein Y19905.02; protein Y1L046w  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 28-May-1993 #sequence\_revision 24-Feb-1995 #text\_change 26-May-2000  
 C:Accession: S49932; S43750  
 R:Odell, C.; Bowman, S.  
 submitted to the EMBL Data Library, December 1994  
 A:Reference number: S49931  
 A:Accession: S49932  
 A:Molecule type: DNA  
 A:Residues: 1-640 <CODE>  
 A:Cross-references: GB:247047; EMBL:246861; NID:6603997; PID:9763300; MIPS:Y1L046w  
 R:Thomas, D.; Cherest, H.; Barbey, R.; Surdin-Kerjan, Y.  
 submitted to the EMBL Data Library, December 1993  
 A:Reference number: S43750  
 A:Accession: S43750  
 A:Molecule type: DNA  
 A:Residues: 1-60; 'I', 62-640 <THO>  
 A:Cross-references: EMBL:L26505; NID:g432493; PID:g432494  
 C:Genetics:

A:Gene: SGD:MET30  
 A:Cross-references: SGD:S0001308; MIPS:Y1L046w  
 A:Map position: 9L  
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
 F:298-339/Domain: WD repeat homology <WD>  
 F:338-369/Domain: WD repeat homology <WD>  
 F:374-409/Domain: WD repeat homology <WD>  
 F:417-450/Domain: WD repeat homology <WD>

Query Match 18.0%: Score 545; DB 2: Length 640;  
 Best Local Similarity 28.9%: Pred. No. 3.2e-34;  
 Matches 167; Conservative 85; Mismatches 202; Indels 124; Gaps 17;

```

QY 34 IPEKNSLRQYNSCARCLN-OETVCLASTAMKTEVCAR-TKLNGTSSMIVYKORRLS 91
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 86 LPEYNTKFCYRHNDDIOPSPHTACQKQDLKRFQEIINAKIPLQESDIHHIISKYS 145
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 92 ASYEKEKELCVYFEOWSESDQVEVEHLISOMCHYGHINSYLKPMLODFTTALPAR 151
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 146 NSNDKIRKL-----ILDGLISTSCFPOLSYSLSLVTHMKIKIDFISLP-- 188
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 152 GLDHTAENILSYLDKSLCAAEVCKEYRTSDGMLKILIERMVRTDSLIRGLAERR- 210
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 189 --QELSLKILSYLDCQSLCNATRCRKWKQKLAADDRVYHMCQHT-----DRKC 236
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 211 ---GWGYLFRNK-----PPDGNAPPNSFYRALYPKIIDTETTESNMRGRLSLQ 258
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 227 PNCWGGLPLHMKRRARIQNSSTGSSNADIQTTIRPKVIYRERFYESWVRG----- 291
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 259 RIHCRSEK-----GVYLODYDQKIVSGLRDNTIKIMKNTLECKRIILTGHTGSLCQ 314
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 292 --HCHIOEFKGMDDVLRLGFNYRLFGSYDSFTIGLWDLPTGLIRLSGHSQGVITLY 349
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 315 YDERVITIGSSDSTVRYVMDVNTGEMLNTLIHCEAVLHLRFNNGMATVCSKRSIAVMD 374
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 350 FDDRKLITLIGSLDKTRVWNYITTEGISTYRGHSBSVLSYDSYQVYVSGADKTVKWHV 409
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 375 ASPDTITLRLVYGHRAAVNVDFDDKYI--VASAGRTIKVWNTSTCEFTVTLNGH--- 429
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 410 ESRTCYTLR---GHEWVNCVYLHPKSFSCFSCSDPTTMMMDIRNMSCLKVRGHVQ 465
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 430 KRGIACLOYRD--RLVSGSS----- 448
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 466 VOKIIPLIKIVENLATDNTSDGSSPODDPTMTGADSDSPNSBOETVLDENIPYTHL 525
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 449 -----DNTIRLMDIFCGACLRVLEGEHELVCIRPDNKRIYSGAYDGIKVMVDVAALDP 503
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 526 LSCGIDNTIKIMDYTGKICIRFGHVGCVMDIADNFRITSGSHDSIKWMDLQSC--- 582
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 504 RAPAGTLCRLTVESHGVRFLQDEFOIVSSSHDIT 541
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 583 -----KCMHFP---NGR--RLQRETHQTOYQSLGDXY 609
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

# RESULT 6

T6660  
 sulfur controller-2 protein [imported] - Neurospora crassa  
 C:Species: Neurospora crassa  
 C:Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 21-Jul-2000  
 C:Accession: T6660  
 R:Kumar, A.; Paletta, J.V.  
 Proc. Natl. Acad. Sci. U.S.A. 92, 3343-3347, 1995  
 A:Title: The sulfur controller-2 negative regulatory gene of Neurospora crassa encode  
 A:Reference number: T6660  
 A:Accession: T6660  
 A>Status: preliminary; translated from GR/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-650 <KUD>  
 A:Cross-references: EMBL:U17251; NID:9806757; PIDN:AAA68968.1; PID:9806758  
 C:Genetics:  
 A:Gene: scon-2  
 A:Map position: 3  
 A:Introns: 75/3; 319/1; 354/1



OY 460 GACIARVLEHHEELVRCIRREDNKRIYSGAYDGKIKYWDVLAALDPRAPAGTLCRLTLVEH- 518  
 DB 180 GKLLRMLGHEHDLIRTVRENDERIVSGYDGTVRIN-----FNTGEQHC 224  
 OY 519 -----SCRVRRLQFDEFOIVSSSHDDTLILWDF 546  
 DB 225 VLHNSRNSRVFGIQPHRRITIACTHSELVWNE 258

## RESULT 9

T18521

beta transducin-like protein - Podospora anserina

C:Species: Podospora anserina

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T18521

R:Saupé, S.; Turcq, B.; Bequeret, J.

Gene 162, 135-139, 1995

A:Title: A gene responsible for vegetative incompatibility in the fungus Podospora anser

A:Reference number: Z18944; MUID:9600981; PMID:7557402

A:Accession: T18521

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1356 &lt;SAU&gt;

A:Cross-references: EMBL:L28125; NID:g607002; PID:g607003; PIDN:AAA85775.1

C:Genetics:

A:Gene: het-el

A:Introns: 761/3

Query Match 15.0%; Score 455.5; DB 2; Length 1356;  
 Best Local Similarity 26.0%; Pred. No. 8.2e-27;  
 Matches 164; Conservative 102; Mismatches 220; Indels 145; Gaps 30;

OY 28 EPRKTIIPKNSLRQITNSCAR-ICLNOETVCLASTAMK-----TENCVAKTKLANGTSS 81  
 DB 581 EMPDVSDDPESELEIVKLCGSPILIRKTVYVHOSAKPFLIGTASDKASNSAQDAFE 640  
 OY 82 MIYPK-----QRKLSASYEKEK--LCVYFPEOWSEDOVEFEHL--ISOMCHY 127  
 DB 641 LVFTGTEDVSYITFMRSLSWMSGLKLRDLICLNAPFLDNNRVPPDPLATVRYSCIT 700  
 OY 128 QHGHINSYLRP-----MLQRD-----FTT-----ALPANGLDHI--AEN 159  
 DB 701 WIDLRLPLVSTSSKRWVLLDDDDIRHPLFTLKYWLALSLRLARP--EGINAIRQLES 759  
 OY 160 ILTSLDKSLCAALVCKEWRVTSDDAMKRLIER--WVRTSLRGLAERGMGOYL 216  
 DB 760 LIGHITIRGLIA--IVRDGIRF--ALSTRMLLEKAPLQAYISAL-----V 800  
 OY 217 FKNRPDGNAPNSFYALY-----PKTIODIETIESNMRCGRSL----- 257  
 DB 801 F-----APTSDMIKKIFKEKPGMISTISVEAEINACTQTLLEGHSSVLSVAFSAD 852  
 OY 258 -QRHCHNS-----ETSKG-----YCCLOY--DDCKIYSGLRDNTIKIMDK 294  
 DB 853 GORVASGSDDTIKIMDTASGTQTLLEGHGSVSWAFSPDRERVASGSDTKTIKIMDA 912  
 OY 295 NTLCKRLITGHTGSVLCLOY--DERVITGSSDSTVAVMDVNTGEMNTLIHHCYAVLH 352  
 DB 913 ASGCTQTLLEGHGRVQSVAFSPDQORASSSDHTIKIMDAASGTCTQTLLEGHSSVLS 972  
 OY 353 LRFN-NGMMVTC-SKDRSIAVMDASPTDITLRVLYVGHRAAVNVDF--DDKIYVASC 408  
 DB 973 VAFSPDQORVASGSDTKIKIMDTASG--TCTQTLLEGHGSVSWAFSPDQORVASGSD 1029  
 OY 409 DRTIKVMTSTCEFRVRLTNGHKRGITACLOYR--DRLVYVSGSDMTIRLMDIECGACLYL 466  
 DB 1030 DRTIKIMDTASGTCTQTLLEGHGRVQSVAFSPDQORASSGSDHTIKIMDAVSGCTQTL 1069  
 OY 467 EGHEELVRCIRF--DNKRIVSGAYDGKIKYWDVLAALDPRAPAGTLCRLTLVHSGRYER 524  
 DB 1090 EGHGDSVSWAFSPDQORVASGSDGTIKIMD-----AASGT--CTQTLLEGHGRVHNS 1140

OY 525 LQF--DEFOIVSSSHDDTLILWDFLNDPAAQ 553  
 DB 1141 VAFSPDQORVASGSDGTIKIMDAASGTCTQ 1171

## RESULT 10

T45136

WD repeat protein pop1 [imported] - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 18-Feb-2000

C:Accession: T45136; T40157

R:Kominami, K.; Toda, T.

submitted to the EMBL Data Library, September 1996

A:Description: Fission yeast WD repeat protein Pop1 is involved in maintenance of plo

A:Reference number: Z22925

A:Accession: T45136

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-775 &lt;KOM&gt;

A:Cross-references: EMBL:Y08391; PIDN:CAA69671.1

A:Experimental source: strain h-972

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Devlin, K.; Churcher, C.M.

submitted to the EMBL Data Library, March 1998

A:Reference number: Z21842

A:Accession: T40157

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-775 &lt;WOO&gt;

A:Cross-references: EMBL:AL022103; PIDN:CAA17898.1; GSPDB:GN00067; SPDB:SPBC262.18

A:Experimental source: strain 972h; cosmid c262

C:Genetics:

A:Gene: SPBC262.18

A:Map position: 2

Query Match 14.9%; Score 453; DB 2; Length 775;  
 Best Local Similarity 25.8%; Pred. No. 5.8e-27;  
 Matches 162; Conservative 86; Mismatches 229; Indels 150; Gaps 20;

OY 9 QEKALFNMSSEREDC-----NNGEPRKTIIPKNSLRQTY--NSCARLCLNOET 56  
 DB 121 REKCLKRRNSLSLNNHAKRFLFNQSDGNKKNEFFPSTNYSNVFPNCDCKEVAEST 180  
 OY 57 VC-----LASTAMKTENCVAKTKLANGTSSMIYPKOKKLSA----- 92  
 DB 181 TFSLDAPNNNSVNTSYSPNMLGNDSTKTRQSPFPHSSSSHNSLHEPVIYDFSSNPSIHP 240  
 OY 93 -----SYEKEKEEL--CVKYFPEOWSEDOVEFEHLISOMCHYQHGHINSYLRKPLQDRF 144  
 DB 241 SNHLSOKNAVLAOLISSFEKLPESVROYILFHLISCKGKHAVONHKKILLPIFOKNF 300  
 OY 145 ITALPARGLDHIENILSYLDAKSLCAELVCKEWR--VTSQGMIMKTL-----I 193  
 DB 301 LTGFPA-----ETVNLVTLHDAPSLCAVSGVSHHWKLVSSNBEMLKSLFLKDGCFWDSI 356  
 OY 194 ERVAVRDSILMRGLAERGMGOYLFKMKPRDGNAPNSFYALYPKTIODIETIESNWRG 253  
 DB 357 DSKIRFMCLOSLSA-----CALMKRYFFHFNIREW----- 389  
 OY 254 RSHLQRLHCHSEFSKGYVC--LOYDDOKIVSGLRDNTIKIMDKNTLECKRLITGHTGSV 310  
 DB 390 LHAPEKIKRGSFPIHGRRLITKQFPDDKLTIVTSGPRINIVDTKGVILRSLIEHEGCV 449  
 OY 311 LCIQYDERVITGSSDSTVAVMDVNTGEMNTLIHHCYAVLH-REFNNGMMVTCSSKRSTA 370  
 DB 450 WTEFEYDVLTVTGSTDTRTVAVMDLRTGCEKQVEYGHSTIRCIKIVQGNOSTDITD- 505  
 OY 371 VMDASPTDITLRVLYVGHRAAVNVDFDDKIYVASCGRDTIKVMN----- 416  
 DB 506 -----DVEKE-----NRPASNDANSMPYIISRRDCTILMSLPCLDDPPFYVNV 552  
 OY 417 -----TSTCEFRVRLTNGHK--RGITACLOYRDLVYVSGSDMTIRLMDIECGAC 462

Db 553 NPDQNDFTSATNPDEYIRLIRGHTDSRVFACL--GDLIVASXDTLRWAKASTGVC 609  
 Oy 463 LRVLEGHELYRCIRFPNKR--IYSGAYDGKIKVMDIYALDPRAPGTLCTRTVHSG 520  
 Db 610 LHVLRGHWGRVYSTVITPNSRQOCISAGTDAKIRIMNL-----BSGEL-IQTLHGSHN 660  
 Oy 521 RVFRLOQDEFOIVSS--HDDTILIMD 545  
 Db 661 LVSGVTFNQNLIVASASAPPTSLRVM 687

## RESULT 11

MD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)  
 C:Species: Nostoc sp.  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
 C:Accession: AE1810  
 R:Kanehiko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete genomic sequence of the filamentous Nitrogen-fixing Cyanobacterium Anabaena  
 A:Reference number: AB1807; MIMD:21595285; PMID:11759840  
 A:Accession: AE1810  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1227 <KUR>  
 A:Cross-references: GB:BA000019; PIDN:BA077553.1; PID:417135007; GSPDB:GN00179  
 A:Experimental source: strain PCC 7120  
 C:Genetics: alr0029

Query Match 13.6%; Score 413.5; DB 2; Length 1227;  
 Best Local Similarity 23.9%; Pred. NO. 1.3e-23;  
 Matches 140; Conservative 114; Mismatches 208; Indels 125; Gaps 22;

Oy 10 EKALKPMNSERE-----DCNNGEPKRIITEKNSLAKOTYNSCARLINDFTVCLA 60  
 Db 378 EKVYVMSLGNKRYYSFRELKDDMLTTESPIKVMLESLR-----RSLEIKASPLLI 431  
 Oy 61 STAKTENCVAKTKLANGTSSMIVPQKRLASAYEKELCVKYPEQMSDOYEFVHLL 120  
 Db 432 EKASSYQGEKEASSKFGLESYVM--EYITAKFIENS-----LEFSQKTKLDF---- 478  
 Oy 121 ISCMYCHYGHINSTYKPMLODFITLAPARGLDHIAENILSYDAKSLCAAEIVCREWY 180  
 Db 479 -----INTY--PLMK-----ARSLDYI----- 493  
 Oy 181 RYVSDMLMKKLIEMVVRFDLSMRGLARRGQVLFKNKPPDGNAPPN--SFYRALYRK 238  
 Db 494 RQIQRLLIEPYKOKLNTFGELEHLRRMLGTLOKEPLFKKGTAAGNLINLLRQLOLD 553  
 Oy 239 ITQDIETISN-----WRCGRHSI--QRIHCRS-----ETSKGVYCLQY--D 276  
 Db 554 KLPDPSPIIDLSGRDPSGLTIWQAYFKEVYKLEKTIFFANSDLTGVSFTLEMSSVVSVKRSPD 613  
 Oy 277 DQKIVSGLRDNTIKIMDKNTLECKRIILGHGVSVCLOY--DEVYITIGSSDSTVRWVY 334  
 Db 614 GYRPAFGTGLMNGEIRLMQTSNKKQRLITKGTAWMAFAFSPDSMLASGSADSTIKIMDV 673  
 Oy 335 NTGEMINLTIHRCVAVLHARF--NNGMVAVTSKSRSTIAVMDMASPTDITLRVLYVGRRA 392  
 Db 674 HNGECLTKLSKNTKNSVAVSPGRILASASQOTITKINDIANG--NCOQTILGHDDW 730  
 Oy 393 VAVYDF-----DDK--YIYASAGDRTIKYVNTSTCEVYRTLNGHRGIALQYR--DELVY 444  
 Db 731 VWSYTFSPVTDRLPLLAASSADQHIKIMDVATGKCLTKGHTREHVSFSPDQTLA 790  
 Oy 445 SSSSDNTIRLMDIEGACLRVLEGHELYRCIRF--DNKRIVSAYDGKIKVMDIYALD 502  
 Db 791 SSGEDSTYRLMDVKTGCQWQJTFEHSKRYIVSVRSPDQTLASGGERSTIKIMDIORG-- 848  
 Oy 503 PRAAPAGTLCRLIVHSGRVRLQF--DEFOIVSSSHDDTILIMDFL 547

Db 849 -----ECVWTLMGHSSQVWAIASFDPGRITLISCSDDQTARLMDVI 888

## RESULT 12

cell division control protein CDC4 - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein YPL009W  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 02-Sep-1995 #sequence\_revision 12-Apr-1996 #text\_change 26-May-2000  
 C:Accession: S56245; S48310; A26867; S62304  
 R:Murakami, Y.; Maitou, M.; Hagihara, H.; Shibata, T.; Ozawa, M.; Sasamura, S.I.; Sas  
 submitted to the EMBL Data Library, May 1995  
 A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces  
 A:Reference number: S56186  
 A:Accession: S56245  
 A:Molecule type: DNA  
 A:Residues: 1-779 <MUR>  
 A:Cross-references: EMBL:D50617; NID:g836685; PIDN:BA09229.1; PID:d1009870; PID:g836  
 R:Churcher, C.  
 submitted to the EMBL Data Library, September 1994  
 A:Reference number: S48310  
 A:Accession: S48310  
 A:Molecule type: DNA  
 A:Residues: 1-579 <CHD>  
 A:Cross-references: EMBL:246255; NID:g559925; PIDN:CAA86341.1; PID:g559926; MIPS:YPL0  
 R:Tochem, U.; Byers, B.  
 J. Mol. Biol. 195, 233-245, 1987  
 A:Title: Structural comparison of the yeast cell division cycle gene CDC4 and a relat  
 A:Reference number: A26867; MUID:88011240; PMID:3309335  
 A:Accession: A26867  
 A:Molecule type: DNA  
 A:Residues: 1-459, 'E', 461-779 <YOC>  
 A:Cross-references: EMBL:X05625; NID:g3502; PIDN:CAA29113.1; PID:g3503  
 R:Murakami, Y.  
 submitted to the EMBL Data Library, April 1994  
 A:Reference number: S62302  
 A:Accession: S62304  
 A:Molecule type: DNA  
 A:Residues: 1-779 <MUR>  
 A:Cross-references: EMBL:D31600; NID:g836814; PIDN:BA06495.1; PID:d1007066; PID:g836  
 C:Identities:  
 A:Gene: SGD: CDC4  
 A:Cross-references: SGD:S0001885; MIPS:YPL009W  
 A:Map position: 6L  
 C:Function:  
 A:Description: Initiation of DNA replication; separation of the spindle pole bodies t  
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
 C:Keywords: cell cycle control  
 F:459-494/Domain: WD repeat homology <WD1>  
 F:628-659/Domain: WD repeat homology <WD2>

Query Match 13.2%; Score 399; DB 2; Length 779;  
 Best Local Similarity 24.3%; Pred. NO. 9.1e-23;  
 Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

Oy 56 IYVCLSTAKTENCVAKTKLANGTSSMIVPQKRLASAYEK-----ELCVKTFEOM 108  
 Db 194 TTPLAKTITINN-----NNNADLIESKDSITISEYSDIFSAINNINLPHAVFK-- 244  
 Oy 109 SESDOVEFEVHLISOMCHYGHINSTYKPMLODFITLAPARGLDHIAENILSYDAS 168  
 Db 245 -----NLFRLVANMRSELSDLGITIKONLKRDLITSLP-----EISLITPNYIQDF 294  
 Oy 169 LCAAEIVCKEYR-VTSDGMLMKLI--EMVVRFDLSMRGLARRGQVLFKNKPPDGN 225  
 Db 295 IINSLGVSONNMKTIKSTSLMKRLISENFV-----SPKGF 331  
 Oy 226 APPNSFYRALYKTIQD-----IEF--IESNMRCGRSLORHCRSFTSKGVYCLQYD 276  
 Db 332 NSLNLKLSQKYPKLSQGDRLRLSLLENIFLKNWYKPKYVQPTITLKGMTSVITCLQYE 391  
 Oy 277 DQKIVSGLRDNTIKIMDKNTLECKRIILGHGVSVCLOYDE-RVITIGSSDSTVRWVYN 335

DB 392 DNVTITGADKKMIRYDSINKKFLQLOLSDHGDVWALKAHGGIILVSGSTDRTRVMDIK 451  
QY 336 TGEMLNTLIHCEAVLHLRFNNGMWTCCKRSIAVMDASPTDITLRVLVGHRAAVNV 395  
DB 452 KG-----CCT-----HVFQGHNSTVAC 468  
QY 396 VDEPD-----KIYASASGDTIKVWMT-----STGE-----PVRTLNG 428  
DB 469 LDIVCEKNIKRYITVGRSDTLTHWKLPEKESVDPHGEEDHPLVFETTPPENPFVGVLRG 528  
QY 429 HKRGACIQRLRYVSGSSDNTIRLMDIEGCACLVRLVGHHELVACIFEDN--KRIYVG 486  
DB 529 HMAVTVVSGRGNIYVSGSYDNTLIWDAQMKCLTILSGHTRITYSTYDHEKRCJISA 588  
QY 487 AYDGKIRVWL-----VALDPRAPAGTL--CLRTLVHSGRVFRLQDFEFOIYSS 535  
DB 589 SMDVTIRIMDLNINWNGESVATNSAPCAKILGAMVTLQGHVTLVGLRLSDKFLVSA 648  
QY 536 SHDITLILDFLMDPAQAAPPSPSRITTY 566  
DB 649 AADGISRGWD-AND-----YSRKFSY 668

## RESULT 13

F-box/WD-repeat protein pop2 - fission yeast (Schizosaccharomyces pombe)

M.Alternate names: proteolysis factor sud1p

C.Species: Schizosaccharomyces pombe

C.Date: 21-Jan-2000 #sequence-revision 21-Jan-2000 #text-change 21-Jul-2000

C.Accession: T43557; T38794; T43798

R.Wolf, D.A.; Jackson, P.K.

submitted to the EMBL Data Library, December 1997

A.Description: Fission yeast pop2 encodes a novel F-box/WD-repeat protein involved in th

A.Reference number: Z22576

A.Accession: T43557

A>Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: mRNA

A.Residues: 1-703 <MO>

A.Cross-references: EMBL:AF038867; PIDN:AAB95480.1

A.Experimental source: strain h-972

R.Gentles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, August 1996

A.Reference number: Z21812

A.Accession: T38794

A>Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: DNA

A.Residues: 1-703 <GEN>

A.Cross-references: EMBL:Z98602; PIDN:CA811275.1; GSPDB:GN00066; SPDB:SPACAD7.03

A.Experimental source: strain 972h-; cosmid c4D7

R.Jallepalli, P.V.; Tien, D.; Kelly, T.J.

Proc. Natl. Acad. Sci. U.S.A. 95, 8159-8164, 1998

A>Title: Sud1 targets cyclin-dependent kinase-phosphorylated Cdc18 and Runt proteins fo

A.Reference number: Z22686; MUID:98318628; PMID:9653157

A.Accession: T43798

A>Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: DNA

A.Residues: 1-703 <CAL>

A.Cross-references: EMBL:AF064515; NID:g3293382; PIDN:AAC39496.1; PID:g3293383

C.Genetics:

A:Gene: pop2; sud1; SPACAD7.03

A:Map position: 1

C:Function:

A:Description: regulated to prevent spontaneous re-replication

Query Match 13.1%; Score 396.5; DB 2; Length 703;  
Best Local Similarity 23.3%; Pred. No. 1.2e-22;

Matches 146; Conservative 92; Mismatches 226; Indels 163; Gaps 25;

QY 14 KFNSSERED--CNGEPPKTI-----IPEKNSLRQTYNSCARLCT---NDETVCL 59  
DB 96 KFNSEIENSYKCLDHPSPVPGDPSISLVQRNLF-YSHSLSPKTIISIDRNNIKID 154

QY 60 ASTAKTENCYAKTKLNGTSSMIVPQKOKLSATYK EKLCKYTEQWSESQVEFVE 118  
DB 155 NSISSNSDNPSPPKV--DTSNTVSPGSKPISELEDLNTQSIQVTEDEPEQIO-SYAF 211  
QY 119 HLISOMCHYOHGH-INSYLRKPLQDFITLAPAGLDHIAINIISYDAKSLCAAEVCK 177  
DB 212 FOLLRSCNRQSMRLLNCEPELKKDILSNLP-----SIYSLINDLHSLCRLVSP 267  
QY 178 EMYRV-----TSDGMFLMKLIERVNTDSLWRGLAERRGMYLFRKNPPDGNAPP-----N 229  
DB 268 TWNRILDVHTS---YKIMF-----SLFGQINENMW-----KYANPNLRPFPLAND 312  
QY 230 SFYRALYRKTIQDIQETTESNRCGHSIOR----- 259  
DB 313 QISDDYFPEIRK-----RHFLNRRKWLFPSPISHLSPFIHPNMTISLLH 360  
QY 260 -----HCRSETSK-GVYCYLDQDKIVSGIRNDITKMDNT 296  
DB 361 KDRITITGSGTQIHNAITGVLEARLKGKEGVAVAKIHEMTLVSGSIDTVVWNIER 420  
QY 297 LECKRILTGHTGSVLCIOY-----DERVITGSSDSTVRVWNTGEML 340  
DB 421 AKCTHIFRGHISILRCLELIVPSRLIRHGVETLVEPDQYIVSGSRDHTLRW----- 472  
QY 341 NTLHHCFAVLHFRNNGMAYTCKRNSIAVMDASPTDITLRVLYVGHRAAVVDPDD 400  
DB 473 -----KLPKNTDPPYLPDNTNSIDRWE-KNPFV--HTLIGHTSVRTISGYG 517  
QY 401 KIYASAGDRITKVNSTSCFVPTLNGHKGICLQYRD--RLVSGSSNTRTIDIE 458  
DB 518 DILVSGSDSIRIRVSTGECLYHGRHSIRITSVLYEPPRNICISGSMKSTRVMDLS 577  
QY 459 CGACLRVLEGEHELRCIRFDPNKRIVSGAYDGKIKVMDVLAALDPRAPAGTLCTRYEH 518  
DB 578 TGTCKVLEGHDAFVTLTNVFNRLISGSADSTIRIMD-----LMTGKR-----LMTVPSN 628  
QY 519 SGRVRLQDFEQIVSSSHDITILIMD 545  
DB 629 SCYISFVSDEHKIT-SGNDGSVKLMD 654

## RESULT 14

WD-repeat protein [imported] - Nostoc sp. (strain PCC 7120) plasmid PCC7120alpha

C.Species: Nostoc sp.

A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C.Date: 14-Dec-2001 #sequence-revision 14-Dec-2001 #text-change 30-Jun-2002

C.Accession: A12493

R.Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata

DNA Res. 8, 205-213, 2001

A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A.Reference number: AB1807; MUID:2159285; PMID:11759840

A.Accession: A12493

A>Status: preliminary

A.Molecule type: DNA

A.Residues: 1-1189 <KUR>

A.Cross-references: GB:BA000020; PIDN:BA878213.1; PID:g17135667; GSPDB:GN00180

A.Experimental source: strain PCC 7120

C.Genetics:

A:Gene: alt7129

A:Genome: plasmid

Query Match 12.5%; Score 378.5; DB 2; Length 1189;  
Best Local Similarity 26.7%; Pred. No. 6.5e-21;  
Matches 96; Conservative 63; Mismatches 107; Indels 93; Gaps 10;

QY 280 IYSGLRNDITIKMDKNTLECKRILTGHTGSVLCIOYD--ERVITGSSDSTVRVWNTG 337  
DB 745 IASGSFDONVKKMLDHTKCVMTLQGHGVYTVSAFNPCKMLLSGSYDQSVKWDKRTG 804  
QY 338 EMLNLTLLHCEAVLHLRFN----- 356

Db 805 RCLDTLKTRNIRWSVAFHPOGHLFVSGDDHAAKIWEIETGQCICKTEGHSNATYTIH 864  
 QY 357 ---NGMWTCKSKDRSIANWD--MASPTDILR-----RYLVGHRAAVNVDEDD--KIYVS 405  
 Db 865 NWEHSLASGHEDOTIKLMDLNLSPKRSNNTHPERILQSHSNRVSFVSSTGQLLAS 924  
 QY 406 ASGDRITKVMNTSTCEFRVTLNGHKGRIACLOVR--DRLVYSGSDNTIRLMDIEGACL 463  
 Db 925 GSADRTIKWSPHTGQCLHTJHGSWMWATAFSLDDKLASSYDHTVKIMDVSSGQCL 984  
 QY 464 RYLEGHELVACIRF--DNKRIVSGAYDGKIKWDL-----VAALDP 503  
 Db 985 QTIQGHPSYLAFAFSCDGKTLFSSGEKLVKQWDVETGYCLQTEADSNRWMAVAVSRL 1044  
 QY 504 RAPAGT-----LCRLTVEHSGRVFRLOF--DEFOIVSSSHDRTILMD 545  
 Db 1045 NOYLATGDDSDSVRLMDIGKGVCTRTSGHTSOYICILFTKDGRRMISSSSDRTIKIMN 1103

## RESULT 15

AC1842

MD-40 repeat protein (imported) - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002

C:Accession: AC1842

R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A:title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AC1842

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1747 &lt;CUR&gt;

A:Cross-references: GB:BA000019; PIDN:BAF7807.1; PID:q17135261; GSPDB:GNO0179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all0283

Query Match

Best Local Similarity 32.7%; Pred. No.2.le-20;

Matches 84; Conservative 50; Mismatches 107; Indels 16; Gaps 7;

QY 276 DDQKIVSGLRDNTIKIMDKNTLECKRIITLGHGTVLCLQY--DERVITIGSSDSSTVAVMD 333  
 Db 1487 DGTIATASADNTIKIMDSQTOQLIKTLTGHKRITTLSPHPDQTIASGSADXTIKIWR 1546  
 QY 334 VNTGEMNTLTHHCEAVLHLEF--NNGMWTCKSKDRSIANWDNASPTDITLRYLVGHR 391  
 Db 1547 VNDGQLRTLTIGHNDEVTYVNSPQGLASGSTDNTVKIWO---TDGRLINKITGHGL 1602  
 QY 392 AVNVVDF--DDKYIVSASGDRITIKVMNTSTCEFRVTLNGHKGRIACLOVR--DRLVYSGS 447  
 Db 1603 AIAVSKFSPDSHTLASASMDNTIKIMQVTDOKLINLNGHIDGVTSLSPDGEILLASGS 1662  
 QY 448 SDNTIRLMDIEGACLVLEGHELVACIRF--DNKRIVSGAYDGKIKWDLVAALDPRA 505  
 Db 1663 ADNTIKIMNPNATLTLKILGHPKINTLAFSPDGKTLISGGEDAGVYVWNL--DLDDL 1720  
 QY 506 PAGTLCRLTVEHSGRV 522  
 Db 1721 QQGGCDRTIDYLOHNSNV 1737

Search completed: February 20, 2003, 09:55:48  
 Job time : 28 secs

Pred. No. is the number of results predicted by chance to have a

ARENZANA SEISEDOS FERNANDO (FR); CONCORDET JEAN PAUL (FR); INST  
NAT SANTE RECH MED (FR); KROHL MATHIAS (FR); DURAND HERVE (FR);  
BENNAOUS RICHARD (FR); MARGOTLIN FLORENCE (FR); PASTEUR INSTITUT  
(FR)

FEATURES  
source

Location/Qualifiers  
1..2151

/organism="synthetic construct"

/db\_xref="taxon:32630"

/note="ADNC codant pour la proteine humaine B7Cp"

70..1779

/note="unnamed protein product"

/codon\_start=1

/transl\_table=11

/protein\_id="CAC07740.1"

/db\_xref="GI:10043428"

/translation="MDPAEAVIQEALKEFMSSEEDCNCEPPRKIIPEKNSIRKQY

NSCARCLNNOETVCLASTAMKTENCVAATKILANGSSITVYKORLISASYKEELCY

KYFEWSESDVEVEHLSIOMCHVQHGHSYLYKPMLODFITALLPARGIDHIAENI

LSYDAKSLCAELVCEKEMRYVSDGIMLKLIERMYVDSLMGLAERGMQGYLTK

NKPDGNAPNSFYKALPKLIODIETTESNMGRSHLSIHRSETSKGYCLOYD

DOKIVSGLRDNTIKIMDKNTLECKRIILGHVLCLOYDERVITITSSDSYVYMDY

MTGEMLNLHHCFAVLIHLRNNGMWYTCSDRSIAYWDASPTDITIRLVYGRHA

VNVYDFDKYIVSAGDPTIKVWMTSTCEPRTNGHKGITACLOIDRLVSSSDN

TIRLMDICGACLVRLVGBHELVICIRPDNKRIVSGAYDKIKWMDLVADLPAPAG

TLCELTIEVSHSFRVRLQFDEFQIVSSHSDTILIMFDLPAAQAEPSPSRITYY

ISR"

BASE COUNT 628 a 467 c 513 g 543 t  
ORIGIN

Query Match 100.0%; Score 2151; DB 6; Length 2151;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCGGTTGCTGCGGCTGGACACCAAGGGGCGCCCGGAGAGAGCGGACCGAGTGGCC 60  
1 TCGGTTGCTGCGGCTGGACACCAAGGGGCGCCCGGAGAGAGCGGACCGAGTGGCC 60  
61 TCGGGGATTATGACCCCGGCGGCGGCTGCTGCAAGAAAGGCACTCAAGTTATGAT 120  
61 TCGGGGATTATGACCCCGGCGGCGGCTGCTGCAAGAAAGGCACTCAAGTTATGAT 120  
61 TCGGGGATTATGACCCCGGCGGCGGCTGCTGCAAGAAAGGCACTCAAGTTATGAT 120  
121 TCGTACAGAGAGAGAGCTGTAATTAATGGGACCCCTAGAGAGATTAATCCAGAGAG 180  
121 TCGTACAGAGAGAGAGCTGTAATTAATGGGACCCCTAGAGAGATTAATCCAGAGAG 180  
181 AATTCACCTTAAGACAGACATACAGACAGCTGCTGCAAGTCTCTAAACCAAGAAACAGTA 240  
181 AATTCACCTTAAGACAGACATACAGACAGCTGCTGCAAGTCTCTAAACCAAGAAACAGTA 240  
241 TGTATAGACAGACCTGCTATGAAGAGTGAATGTGTGGCCAAACAACTTGGCAAT 300  
241 TGTATAGACAGACCTGCTATGAAGAGTGAATGTGTGGCCAAACAACTTGGCAAT 300  
241 TGTATAGACAGACCTGCTATGAAGAGTGAATGTGTGGCCAAACAACTTGGCAAT 300  
301 GGCACCTTCACATGATGTGTGCCAAGCAAGGAAACTCTCAGCAAGCTATGAAAGGAA 360  
301 GGCACCTTCACATGATGTGTGCCAAGCAAGGAAACTCTCAGCAAGCTATGAAAGGAA 360  
301 GGCACCTTCACATGATGTGTGCCAAGCAAGGAAACTCTCAGCAAGCTATGAAAGGAA 360  
361 AAGGAACCTGTGTCAAACTCTTGAAGAGTGTGCAAGTCAAGTCAAGTCAAGTCAAGT 420  
361 AAGGAACCTGTGTCAAACTCTTGAAGAGTGTGCAAGTCAAGTCAAGTCAAGTCAAGT 420  
361 AAGGAACCTGTGTCAAACTCTTGAAGAGTGTGCAAGTCAAGTCAAGTCAAGTCAAGT 420  
421 GAACATCTTATATCCCAATGTGTATTAACCAATGAGGCAATTAACCTGATCTTAAA 480  
421 GAACATCTTATATCCCAATGTGTATTAACCAATGAGGCAATTAACCTGATCTTAAA 480  
421 GAACATCTTATATCCCAATGTGTATTAACCAATGAGGCAATTAACCTGATCTTAAA 480  
481 CCAATGTGTGAGAGATTTCAATGCTGTGCAAGTCTGGGAGATTGATCAATATGCT 540  
481 CCAATGTGTGAGAGATTTCAATGCTGTGCAAGTCTGGGAGATTGATCAATATGCT 540  
541 GAGAACATTTCTGTCAATCTGGAGATGCAATCACTATGTCTGTGAATCTGTGTCAAG 600  
541 GAGAACATTTCTGTCAATCTGGAGATGCAATCACTATGTCTGTGAATCTGTGTCAAG 600

601 GAATGTACGAGAGTACCTCTGATGGATGCTGTGGAAGAGCTTATCGAGAGATGTC 660  
601 GAATGTACGAGAGTACCTCTGATGGATGCTGTGGAAGAGCTTATCGAGAGATGTC 660  
661 AGGACAGATTTCTGTGTGAGAGAGGCTGTGCAAGACAGAGAGATGGGACATTTATTC 720  
661 AGGACAGATTTCTGTGTGAGAGAGGCTGTGCAAGACAGAGAGATGGGACATTTATTC 720  
721 AAAAACAACCTCTGAGAGGAGATGCTGTGCAAGACAGAGAGATTTATAGACATTTATTC 780  
721 AAAAACAACCTCTGAGAGGAGATGCTGTGCAAGACAGAGAGATTTATAGACATTTATTC 780  
721 AAAAACAACCTCTGAGAGGAGATGCTGTGCAAGACAGAGAGATTTATAGACATTTATTC 780  
781 AAAATTTATCAAGACATTTGAGACATTAATTAATTTGAGATGTGGAACATAGTTTA 840  
781 AAAATTTATCAAGACATTTGAGACATTAATTAATTTGAGATGTGGAACATAGTTTA 840  
841 CAGAGATTTCTGCTGCGAGAGTGAAGACAGAGAGATTTATAGACATTTATTC 900  
841 CAGAGATTTCTGCTGCGAGAGTGAAGACAGAGAGATTTATAGACATTTATTC 900  
901 CAGAAATTTAGTAAAGCGGCTTTCGAGACACACATCAAGATCTGGATTAACACATTTG 960  
901 CAGAAATTTAGTAAAGCGGCTTTCGAGACACACATCAAGATCTGGATTAACACATTTG 960  
961 GAATGTACGAGAGATTTCTGAGAGGAGATGCTGTGCAAGACAGAGAGATTTATAGAC 1020  
961 GAATGTACGAGAGATTTCTGAGAGGAGATGCTGTGCAAGACAGAGAGATTTATAGAC 1020  
1021 AGAGTATCAATTAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
1021 AGAGTATCAATTAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
1081 GAATGTAAACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
1081 GAATGTAAACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
1141 GGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
1141 GGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
1201 ACTGACATTAACCTCTGAGAGGAGTGTGCTGCAAGACAGAGAGTCTGATGATGATGAT 1260  
1201 ACTGACATTAACCTCTGAGAGGAGTGTGCTGCAAGACAGAGAGTCTGATGATGATGAT 1260  
1261 TTTGATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
1261 TTTGATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
1321 AGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
1321 AGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
1381 TACAGGAGAGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
1381 TACAGGAGAGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
1441 GAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
1441 GAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
1501 TTTGATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
1501 TTTGATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
1561 GTGGGCTGTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620  
1561 GTGGGCTGTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620  
1621 CATTCCGAGAGATTTTTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
1621 CATTCCGAGAGATTTTTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
1681 GATGACACATCTCTATCTGGAGCTTCTTAATGATCCAGCTGCCAAGCTGAACCCGCC 1740  
1681 GATGACACATCTCTATCTGGAGCTTCTTAATGATCCAGCTGCCAAGCTGAACCCGCC 1740

Db	1681	GATGACACAAATCCATCTGTGGACTTCCTAATATATCATCCAGTCCGCAAGCTGAACCCCC	1740
Qy	1741	CGTTCOCCTTCTGCACTACACCTACATCTCCAGATAAATAAACCATACACTGACCTCAT	1800
Db	1741	CGTTCOCCTTCTGCACTACACCTACATCTCCAGATAAATAAACCATACACTGACCTCAT	1800
Qy	1801	ACTTGCCCGAGNACCATTAAAGTTGCGGATATTTAAAGTATCTGCGCAATACCAGGATAGC	1860
Db	1801	ACTTGCCCGAGNACCATTAAAGTTGCGGATATTTAAAGTATCTGCGCAATACCAGGATAGC	1860
Qy	1861	AACAACAGTAAACATCAAACTACTGCCCCAGTTTCCCTGAGCTAGCCGAGGAGCAGGGCTT	1920
Db	1861	AACAACAGTAAACATCAAACTACTGCCCCAGTTTCCCTGAGCTAGCCGAGGAGCAGGGCTT	1920
Qy	1921	TGAGACTCCTCTTGGGACACAGTTGGTGTGCAAGTGGCCCGAGAGGGTCTACTAGCACA	1980
Db	1921	TGAGACTCCTCTTGGGACACAGTTGGTGTGCAAGTGGCCCGAGAGGGTCTACTAGCACA	1980
Qy	1981	ACTGACTCCTTCAGTGTCTGTATCAGAAAGATGTCTTCATCAATTTGTAATGATTGGAAAC	2040
Db	1981	ACTGACTCCTTCAGTGTCTGTATCAGAAAGATGTCTTCATCAATTTGTAATGATTGGAAAC	2040
Qy	2041	TTTAAACCTCCCTCCTCTCTCTCTCTTCCACTCTGCACCTAGTTTTCCTCATTTGGTTC	2100
Db	2041	TTTAAACCTCCCTCCTCTCTCTCTCTTCCACTCTGCACCTAGTTTTCCTCATTTGGTTC	2100
Qy	2101	CAGACAAGAGTCACTTATAATATATTTAAGTGTTTGGCAGAAAAAANAANA 2151	
Db	2101	CAGACAAGAGTCACTTATAATATATTTAAGTGTTTGGCAGAAAAAANAANA 2151	

[illegible]

Db	181	AATTCACCTTAGACAGACATACAAACAGCTGTGCCAGACTCTGCTTTAAACCAAGAAACAGTA	240
Oy	241	TGTTTACGAAGCACTCTCTATGAAGACTGAGAAATTTGTGGCCAAAACAAACCTTGCCAA	300
Db	241	TGTTTACCAAGCACTCTCTATGAAGACTGAGAAATTTGTGGCCAAAACAAACCTTGCCAA	300
Oy	301	GGCACTTCCAGTATGATGTGTGCCCAAGCAAGGAAACCTCTAGCAAGCTATGAAAGGAA	360
Db	301	GGCACTTCCAGTATGATGTGTGCCCAAGCAAGGAAACCTCTAGCAAGCTATGAAAGGAA	360
Oy	361	AAGGAACTGTGTCTCAAAATACCTTTGAGCAGTGTGACAGTCAAGTCAAGTGAATTTGTG	420
Db	361	AAGGAACTGTGTCTCAAAATACCTTTGAGCAGTGTGACAGTCAAGTCAAGTGAATTTGTG	420
Oy	421	GAAATCTTATATCCCAAAATGTCTATTACCAACATGGGACATTAACCTGTATCTTTAA	480
Db	421	GAAATCTTATATCCCAAAATGTCTATTACCAACATGGGACATTAACCTGTATCTTTAA	480
Oy	481	CCATATGTGACAGAGAGATTTCAATACCTGCTCGCCAGCTCGGGATTTGATCATATTCGT	540
Db	481	CCATATGTGACAGAGAGATTTCAATACCTGCTCGCCAGCTCGGGATTTGATCATATTCGT	540
Oy	541	GAGAAACATTCGTGATACCTGATGGTCCAATACATATGCGTGGTGAACCTTGATGCAAG	600
Db	541	GAGAAACATTCGTGATACCTGATGGTCCAATACATATGCGTGGTGAACCTTGATGCAAG	600
Oy	601	GAATGCTACCGAGTGCACCTCTGATGGCATCTGCTGGAAGAAAGCTTTATCGAGACATGGTC	660
Db	601	GAATGCTACCGAGTGCACCTCTGATGGCATCTGCTGGAAGAAAGCTTTATCGAGACATGGTC	660
Oy	661	AGGACAGATTTCTGTGTGGAAAGGCGTGGCAACGAAAGAGATGGGGACGATATTATTTC	720
Db	661	AGGACAGATTTCTGTGTGGAAAGGCGTGGCAACGAAAGAGATGGGGACGATATTATTTC	720
Oy	721	AAAAACAACCTCCTGACGGGAAATGCTCCTCCCAACTCTTTTATAGACACTTTATCCT	780
Db	721	AAAAACAACCTCCTGACGGGAAATGCTCCTCCCAACTCTTTTATAGACACTTTATCCT	780
Oy	781	AAAAATTATCAAGACATTTGAGACATTAAGATTTCAATTTGAGATGTGGAAGCAATAGTTTA	840
Db	781	AAAAATTATCAAGACATTTGAGACATTAAGATTTCAATTTGAGATGTGGAAGCAATAGTTTA	840
Oy	841	CAGAGAACTTCACTGCGCAAGGAAACAAACGAAAGAGTTTACTGTTTACAGTATGATGAT	900
Db	841	CAGAGAACTTCACTGCGCAAGGAAACAAACGAAAGAGTTTACTGTTTACAGTATGATGAT	900
Oy	901	CAGAAAATAGTAGAGCGGCTTCGAGACACACATCAAGATCGGAGTAAAAACACATTCG	960
Db	901	CAGAAAATAGTAGAGCGGCTTCGAGACACACATCAAGATCGGAGTAAAAACACATTCG	960
Oy	961	GAATGCAAGCGAATTTCTACAGGCGCATACAGGTTCAAGTCTCTGTCTCCAGATGATGAG	1020
Db	961	GAATGCAAGCGAATTTCTACAGGCGCATACAGGTTCAAGTCTCTGTCTCCAGATGATGAG	1020
Oy	1021	AGAGTGATCATTAACAGATCATCGGATTCAGCGGTGAGAGTGTGGATGTAAATACAGT	1080
Db	1021	AGAGTGATCATTAACAGATCATCGGATTCAGCGGTGAGAGTGTGGATGTAAATACAGT	1080
Oy	1081	GAAATGCTTAAACAGCTGATTTACCATTTGGAAGCAATTTGTGACTTGTGCTTTCAATAT	1140
Db	1081	GAAATGCTTAAACAGCTGATTTACCATTTGGAAGCAATTTGTGACTTGTGCTTTCAATAT	1140
Oy	1141	GGCATATGATGATCCTGCTCCAAAGATGCTTCCATTGCTGTATGGATATGGCTCCCA	1200
Db	1141	GGCATATGATGATCCTGCTCCAAAGATGCTTCCATTGCTGTATGGATATGGCTCCCA	1200
Oy	1201	ACTGACATTACCTCTCGGAGGGTGTGTGTGGACACCGAGCTCTGTCAATTTGTAGAC	1260
Db	1201	ACTGACATTACCTCTCGGAGGGTGTGTGTGGACACCGAGCTCTGTCAATTTGTAGAC	1260
Oy	1261	TTTGTAGCAAGATGATTTGTTTCTTCATCTGGGGATAGAACTTTAAAGGTATGAAACACA	1320
Db	1261	TTTGTAGCAAGATGATTTGTTTCTTCATCTGGGGATAGAACTTTAAAGGTATGAAACACA	1320

QY 1321 AGTACTGTGATTTGTAAGACCTTAATATGACACAAAGAGCATTCCTCTTTCAG 1380  
 |||||||  
 Db 1321 AGTACTGTGATTTGTAAGACCTTAATATGACACAAAGAGCATTCCTCTTTCAG 1380  
 QY 1381 TACAGGACAGAGGCTGTAGTGTGCTCATCTGACACACTCTGATTTAGGACATA 1440  
 |||||||  
 Db 1381 TACAGGACAGAGGCTGTAGTGTGCTCATCTGACACACTCTGATTTAGGACATA 1440  
 QY 1441 GATGTGTGCTGATTTAGGACGCTTGAAGGCGCATGAGCAATGGTGGCTGTATTCGA 1500  
 |||||||  
 Db 1441 GATGTGTGCTGATTTAGGACGCTTGAAGGCGCATGAGCAATGGTGGCTGTATTCGA 1500  
 QY 1501 TTTGATTAACAGAGATAGTATGATGAGGCGCTATGATGAAAAATTAAGTGGGATCT 1560  
 |||||||  
 Db 1501 TTTGATTAACAGAGATAGTATGATGAGGCGCTATGATGAAAAATTAAGTGGGATCT 1560  
 QY 1561 GTGGCTGCTTTGGACCCCGCTGCTGACAGGACACTGTCTACAGACCTTGTGAG 1620  
 |||||||  
 Db 1561 GTGGCTGCTTTGGACCCCGCTGCTGACAGGACACTGTCTACAGACCTTGTGAG 1620  
 QY 1621 CATTCGGAAGAGTTTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
 |||||||  
 Db 1621 CATTCGGAAGAGTTTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
 QY 1681 GATGACACATCTCATCTGAGACTTCTTAATGATCCAGTGCACCAAGCTGAACCCGCC 1740  
 |||||||  
 Db 1681 GATGACACATCTCATCTGAGACTTCTTAATGATCCAGTGCACCAAGCTGAACCCGCC 1740  
 QY 1741 GCTTCCCTCTTTCGACACATACATCTCTGACATTAATTAATTAATTAATTAATTAAT 1800  
 |||||||  
 Db 1741 GCTTCCCTCTTTCGACACATACATCTCTGACATTAATTAATTAATTAATTAATTAAT 1800  
 QY 1801 ACTTCCCGACGACCACTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1860  
 |||||||  
 Db 1801 ACTTCCCGACGACCACTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1860  
 QY 1861 ACTTCCCGACGACCACTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
 |||||||  
 Db 1861 ACTTCCCGACGACCACTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
 QY 1921 TGAGACTCTCTGTTGGACACAGTTGTTGCTGACAGTGGGCGGACGAGCTCTACAGACA 1980  
 |||||||  
 Db 1921 TGAGACTCTCTGTTGGACACAGTTGTTGCTGACAGTGGGCGGACGAGCTCTACAGACA 1980  
 QY 1981 ACTGACTCTCTGTTGCTGATCAGAGATGCTTCTATCAATTTGGAATTTGGAC 2040  
 |||||||  
 Db 1981 ACTGACTCTCTGTTGCTGATCAGAGATGCTTCTATCAATTTGGAATTTGGAC 2040  
 QY 2041 TTTTAAACCTCCCT 2100  
 |||||||  
 Db 2041 TTTTAAACCTCCCT 2100  
 QY 2101 CAGACAAAGGAGCTTATTAATATATATATATATATATATATATATATATATATATAT 2151  
 |||||||  
 Db 2101 CAGACAAAGGAGCTTATTAATATATATATATATATATATATATATATATATATATAT 2151

RESULT 3  
 AF129530 2151 bp mRNA linear PRI 31-OCT-1999  
 LOCUS AF129530 Homo sapiens chromosome 10 F-box protein Fbw1A (FBW1A) mRNA,  
 DEFINITION complete cds.  
 ACCESSION AF129530  
 VERSION AF129530.1 GI:6164609  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 2151)  
 AUTHORS Cenciarelli,C., Chlaue,D.S., Guardavaccaro,D., Parks,W., Vidal,M.  
 and Pagano,M.

TITLE Identification of a family of human F-box proteins  
 JOURNAL Curr. Biol. 9 (20), 1177-1179 (1999)  
 MEDLINE 20003060  
 PUBMED 10531035  
 REFERENCE 2 (bases 1 to 2151)  
 AUTHORS Chlaue,D.S. and Pagano,M.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-FEB-1999) Pathology, NYU Medical Center, 550 First  
 Ave. MSB 548, New York, NY 10016, USA  
 FEATURES  
 source  
 1..2151  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="10"  
 /map="10q24"  
 1..2151  
 /gene="FBW1A"  
 70..1779  
 /gene="FBW1A"  
 /function="regulates the stability of IkappaB alpha- and  
 beta-catenin"  
 /note="contains one F-box domain and seven WD40 domains;  
 forms an SCF with Skp1 and Cull1"  
 /codon\_start=1  
 /product="F-box protein Fbw1A"  
 /protein\_id="AF04464.1"  
 /db\_xref="GI:6164610"  
 /translation="MDPAEAVLQEKALPKPNSEREDNNGEPPKIIPEKNSLROTY  
 NSCARCLINDEYCLASTAKTEKNCVAKKILANGNSMIVPORKLSIYEKELCY  
 KYEOWESDOYEFVPHLISOMCHQOHGINSYLPMLORDTIALPARGDLHLENT  
 LSTLDKSLCAEILYKERYRYSQMLKILKEMVRDLSRLARSGVGLFK  
 NRPDGNAPPNSFTALYPIKIIIDIEFTESNRCGRHSIORIHRSETSKGYCLQYD  
 DKIVSGLDNTIKLMDKNTLECKRLIHTGHSVLCLOYDERVITTSDSSTVRAWDY  
 NQEMNLNLIHICEAVLIHLRFNNGMAVCSKDRSLAVMDASPTDITLRYVSGRBA  
 VNVVFPDQYIVSASGDRITKYVNTSTCEVFTLNGHRGLACLOYRRLVYSGSSN  
 TIRLMDIECGACTLVEGHELELRIRDNRIYSGADGKIRKWDVLAALDPRPAG  
 TICLRITVHSRVRFLQDFEQIYSSHDDITLIMDFLNDPAQAEPPNSPSRTIY  
 ISR"  
 BASE COUNT 628 a 467 c 513 g 543 t  
 ORIGIN  
 Query Match 100.0%; Score 2151; DB 9; Length 2151;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TGGGTGGCTGCGCGCTGGCACCACAAAGGGCGGCCGGGAGAGCGGACCGAGTGGCC 60  
 |||||||  
 Db 1 TGGGTGGCTGCGCGCTGGCACCACAAAGGGCGGCCGGGAGAGCGGACCGAGTGGCC 60  
 QY 61 TCGGCGATTATGAGCCCGGCGGCGGCGGTGTCAGAAAGAGAGCACTCAAGTTATGANT 120  
 |||||||  
 Db 61 TCGGCGATTATGAGCCCGGCGGCGGCGGTGTCAGAAAGAGAGCACTCAAGTTATGANT 120  
 QY 121 TCCTCAGAGAGAGAGACTGTATATATATATATATATATATATATATATATATATATAT 180  
 |||||||  
 Db 121 TCCTCAGAGAGAGAGACTGTATATATATATATATATATATATATATATATATATATAT 180  
 QY 181 AATTCATTAGACAGACATACAAAGAGCTGTGCAACACCTCTGTTAAACCAAGAAACAGTA 240  
 |||||||  
 Db 181 AATTCATTAGACAGACATACAAAGAGCTGTGCAACACCTCTGTTAAACCAAGAAACAGTA 240  
 QY 241 TGTTTAGCAAGACAGCTGATGAGAGCTGATGAGAGCTGATGAGAGCTGATGAGAGCTGAT 300  
 |||||||  
 Db 241 TGTTTAGCAAGACAGCTGATGAGAGCTGATGAGAGCTGATGAGAGCTGATGAGAGCTGAT 300  
 QY 301 GGCACCTCCAGATATATGTTGTCGCAAGCAAGAAAGCTGACGAAGCTATGAAAGGAA 360  
 |||||||  
 Db 301 GGCACCTCCAGATATATGTTGTCGCAAGCAAGAAAGCTGACGAAGCTATGAAAGGAA 360  
 QY 361 AAGGAACCTGTGTCAAAATTAATTTGAGCAGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 420  
 |||||||  
 Db 361 AAGGAACCTGTGTCAAAATTAATTTGAGCAGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 420

QY 421 GAACATCTATATCCAAATGTGTCATATACCAATGGGACATATACTGTATCTTAA 480  
 DB 421 GAACATCTATATCCAAATGTGTCATATACCAATGGGACATATACTGTATCTTAA 480  
 QY 481 CCTATGTCAGAGAGATTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
 DB 481 CCTATGTCAGAGAGATTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
 QY 541 GAGAACATCTGTCATACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 DB 541 GAGAACATCTGTCATACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 QY 601 GAATGTACCGAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 DB 601 GAATGTACCGAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 QY 661 AGGACAGATTTCTGTCGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
 DB 661 AGGACAGATTTCTGTCGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
 QY 721 AAAAACAACCTCTGAGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
 DB 721 AAAAACAACCTCTGAGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
 QY 781 AAAATATACAGACATTTGACATATGATCTATATTGAGATGAGAGACATATTTA 840  
 DB 781 AAAATATACAGACATTTGACATATGATCTATATTGAGATGAGAGACATATTTA 840  
 QY 841 CAGAGAAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 DB 841 CAGAGAAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 QY 901 CAGAGAAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
 DB 901 CAGAGAAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
 QY 961 GAATGCAAGGAAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
 DB 961 GAATGCAAGGAAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
 QY 1021 AGAGTATCATATACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
 DB 1021 AGAGTATCATATACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
 QY 1081 GAAATGCTAAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
 DB 1081 GAAATGCTAAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
 QY 1141 GGCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
 DB 1141 GGCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
 QY 1201 ACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 DB 1201 ACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 QY 1261 TTTGATGCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
 DB 1261 TTTGATGCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
 QY 1321 AGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
 DB 1321 AGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
 QY 1381 TACAGGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
 DB 1381 TACAGGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
 QY 1441 GAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
 DB 1441 GAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
 QY 1501 TTTGATGCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560

DB 1501 TTTGATGCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560  
 QY 1561 GTGGCGCTTTGGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
 DB 1561 GTGGCGCTTTGGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
 QY 1621 CATTCGGAAGATTTTGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
 DB 1621 CATTCGGAAGATTTTGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
 QY 1681 GATGACAAATCTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
 DB 1681 GATGACAAATCTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
 QY 1741 CGTCCCTCTGCAACATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
 DB 1741 CGTCCCTCTGCAACATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
 QY 1801 ACTTGCCGAGACCCATTAAGTGGGTATTTAACTGCTGCTGCTGCTGCTGCTGCTGCT 1860  
 DB 1801 ACTTGCCGAGACCCATTAAGTGGGTATTTAACTGCTGCTGCTGCTGCTGCTGCTGCT 1860  
 QY 1861 AACACAGTAACTAACTAACTAACTAACTAACTAACTAACTAACTAACTAACTAACT 1920  
 DB 1861 AACACAGTAACTAACTAACTAACTAACTAACTAACTAACTAACTAACTAACTAACT 1920  
 QY 1921 TGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980  
 DB 1921 TGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980  
 QY 1981 ACTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
 DB 1981 ACTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
 QY 2041 TTTTAAACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100  
 DB 2041 TTTTAAACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100  
 QY 2101 CACAGAAAGTGAATTAAT 2151  
 DB 2101 CACAGAAAGTGAATTAAT 2151

RESULT 4  
 HSBTRCP  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

HSBTRCP  
 Homo sapiens mRNA for beta-transducin repeat containing protein.  
 Y14153.1  
 Y14153.1 GI:2995193  
 beta-transducin repeats; beta-TRCP gene; WD repeat.  
 Homo sapiens.  
 Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
 AUTHORS  
 TITLES  
 JOURNAL

JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS  
 TITLES  
 JOURNAL

Mol. Cell 1 (4), 565-574 (1998)  
 98325370  
 960940  
 2 (bases 1 to 2151)  
 Benarous, R.  
 Direct Submission  
 Submitted (03-JUL-1997) R. Benarous, INSERM - I.C.G.M., Laboratoire  
 Interactions Proteiques, CHU Cochin, 24 rue de Fg-St-Jacques, 75014  
 Paris, FRANCE  
 Location/Qualifiers  
 1..2151  
 /organism="Homo sapiens"

```

/db_xref="taxon:9606"
/cell_line="Jurkat T-cell"
/cell_type="T lymphocyte"
/tissue_type="lymphoid"
70..1779
/gene="beta-TRCP"
70..1779
/gene="beta-TRCP"
/codon_start=1
/product="beta-transducin repeats containing protein"
/protein_id="CAAF4572.1"
/db_xref="GI:2995194"
/translation="MDPAEAVLQEKALFENSSEREDCNNGEPKRIIEPKNSLROTY
NSCARLCLNOETVLAETAMKTEVCYKTLANGTSMVLPKORLISASYEKEKELCY
KYFQWSEDOVEFEHLISOMCHYOHGHTNSYLKPLQDPTALPASGELDIANI
LSYDAKSLCAAEVCKEWRVYSDGLMKKILERWRTDLSMRGLAERGMGVLEK
NPPDCAAPNPSTFRLPYKTIQDIEFTESNMRCGRSLQRIHCSESTGYVCLQYD
DKIVSGLRNTIKIMDKNTLECKRLITGHTSVLCLOYDERVITIGSSPSYVWMDV
NTEMLNTLIHCEAVLHFRNNGMYTCKSRSLIYMDMSPDITLRLVGHRAA
VNVVDFDKYIVASGDRITKVNNTSCEVFTLNGHRIACLOYRDLRVSSSDN
TIRLMDIEGACLRVLEHELYRCIPEPKRIYVSGYDQKIKVMDLVLAIDPRAG
TLCRLVHSGRVFRQPEPQIVSSSHDITLINDLPAAQAAPRPSRRTTY
ISR"

```

## repeat\_region

```

259..544
/rpl_family="WD"

```

```

BASE COUNT      628 a 467 c 513 g 543 t
ORIGIN

```

```

Query Match      100.0%; Score 2151; DB 9; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TCGGTGGCTGCGGCGGCGACCAAGGCGGCGCGCGGCGAGAGAGCGAGCCAGTGGCC 60
Db 1 TCGGTGGCTGCGGCGGCGACCAAGGCGGCGCGCGGCGAGAGAGCGAGCCAGTGGCC 60
QY 61 TCGGCGATATGACCCGCGGCGAGGCGGTGCTGCAAGAGAGGCACTCAAGTTTAAAT 120
Db 61 TCGGCGATATGACCCGCGGCGAGGCGGTGCTGCAAGAGAGGCACTCAAGTTTAAAT 120
QY 121 TCCGACAGAGAGAGAGAGCTGATATGAGGAAACCCCTGAGGAAAGATATACAGAGAA 180
Db 121 TCCGACAGAGAGAGAGAGCTGATATGAGGAAACCCCTGAGGAAAGATATACAGAGAA 180
QY 181 AATTCACTTAGACAGACATCAACAGCTGTCCAGACTGTCTTAAACCAAGAAAGATA 240
Db 181 AATTCACTTAGACAGACATCAACAGCTGTCCAGACTGTCTTAAACCAAGAAAGATA 240
QY 241 TGTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 241 TGTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 GGCACCTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 301 GGCACCTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 361 AAGCACTGTCTGTCAATATCTTGAAGCACTGTCTGTCAAGATCAGATCAAGTGAATTTG 420
Db 361 AAGCACTGTCTGTCAATATCTTGAAGCACTGTCTGTCAAGATCAGATCAAGTGAATTTG 420
QY 421 GAACATCTTATATCCCAATGTCTATACCAATGAGGAGCACTAAACCTGATTTTAA 480
Db 421 GAACATCTTATATCCCAATGTCTATACCAATGAGGAGCACTAAACCTGATTTTAA 480
QY 481 CCTATGTTGAGAGAGATTTTCAATGCTGTCTGCAAGTGGGATTTGATATATCGCT 540
Db 481 CCTATGTTGAGAGAGATTTTCAATGCTGTCTGCAAGTGGGATTTGATATATCGCT 540
QY 541 GAGAACATTTGTCATACCTGAGTGGCAATCTATGCTGCTGCAACTGTGTGCAAG 600
Db 541 GAGAACATTTGTCATACCTGAGTGGCAATCTATGCTGCTGCAACTGTGTGCAAG 600

```

```

QY 601 GAATGTTACCGAGTGAACCTCTGTATGSCATGCTGTGGAAGAGCTTATCGAGAGAAATGTC 660
Db 601 GAATGTTACCGAGTGAACCTCTGTATGSCATGCTGTGGAAGAGCTTATCGAGAGAAATGTC 660
QY 661 AGGACAGATTTCTGTGAGAGAGGCGTGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 661 AGGACAGATTTCTGTGAGAGAGGCGTGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 AAAAACAACCTCTGACGGGAAATGCTCTCCCACTCTTTTATAGACACTTTATCT 780
Db 721 AAAAACAACCTCTGACGGGAAATGCTCTCCCACTCTTTTATAGACACTTTATCT 780
QY 781 AAAATATACAGACATTTGAGACATATGATATATGAGATGTTGAGAGACATGTTTA 840
Db 781 AAAATATACAGACATTTGAGACATATGATATATGAGATGTTGAGAGACATGTTTA 840
QY 841 CAGAGATTCACGCGCGAGAGTGAACACAAAGAGAGTTCATGTTTACAGATATGAT 900
Db 841 CAGAGATTCACGCGCGAGAGTGAACACAAAGAGAGTTCATGTTTACAGATATGAT 900
QY 901 CAGAAAATGTAAAGCGGCTTGCAGACAAACACATCAAGATCTGGGATAAAAACATTTG 960
Db 901 CAGAAAATGTAAAGCGGCTTGCAGACAAACACATCAAGATCTGGGATAAAAACATTTG 960
QY 961 GAATGCAAGGAAATTCACAGAGCCATACAGGTTAGTCTCTCTGCTCAGATATGAT 1020
Db 961 GAATGCAAGGAAATTCACAGAGCCATACAGGTTAGTCTCTCTGCTCAGATATGAT 1020
QY 1021 AGAGTATCATTAACAGAGATCATCGGATTCACAGGTCAGAGTGTGGAGATTAATACAGT 1080
Db 1021 AGAGTATCATTAACAGAGATCATCGGATTCACAGGTCAGAGTGTGGAGATTAATACAGT 1080
QY 1081 GAAATGCTTAAACAGCTGTGATTCACCATTTGTAAGAGATTTGCACTTCCGTTTCAATAT 1140
Db 1081 GAAATGCTTAAACAGCTGTGATTCACCATTTGTAAGAGATTTGCACTTCCGTTTCAATAT 1140
QY 1141 GGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Db 1141 GGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 ACTGACATTAACCTCTCGGAGGCTGCTGTGCGACACCGAGCTGTCAATGTTGTAGAC 1260
Db 1201 ACTGACATTAACCTCTCGGAGGCTGCTGTGCGACACCGAGCTGTCAATGTTGTAGAC 1260
QY 1261 TTTGATGACAACTATATGTTCTGATCTGTGGGATTAACATTAAGTATGGAACACA 1320
Db 1261 TTTGATGACAACTATATGTTCTGATCTGTGGGATTAACATTAAGTATGGAACACA 1320
QY 1321 AGTACTTGTGAATTTGTAGAGACCTTAAATGAGACAAAGAGGATGCTGTTTGAG 1380
Db 1321 AGTACTTGTGAATTTGTAGAGACCTTAAATGAGACAAAGAGGATGCTGTTTGAG 1380
QY 1381 TACAGGAGAGGCTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAG 1440
Db 1381 TACAGGAGAGGCTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAG 1440
QY 1441 GAATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1441 GAATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1501 TTTGATTAACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Db 1501 TTTGATTAACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
QY 1561 GTGGCTGCTTGGAGACCCGCTGTGCTGTGAGAGACCTGTCTACAGGACCTTGTGAG 1620
Db 1561 GTGGCTGCTTGGAGACCCGCTGTGCTGTGAGAGACCTGTCTACAGGACCTTGTGAG 1620
QY 1621 CATTCGGAAGAGTTTTCAGATACAGTTTATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1621 CATTCGGAAGAGTTTTCAGATACAGTTTATGATGATGATGATGATGATGATGATGATGAT 1680
QY 1681 GATGACAAATCTCATCTGTGGAGCTTCTTAATATGATCAAGCTGCCAAGCTGAACCCGCC 1740

```

[illegible]

Db 728 AGAATGGTGCAGACATCTCTGTGGAGAGGCTGGCAAGAGAGAGATGGGACAG 787  
 QY 712 TATTTTTCACAAAACAACTCTCTGACGGGAATGCTCTCCCACTCTTTTATAGACA 771  
 Db 788 TATTTTTCACAAAACAACTCTCTGACGGGAATGCTCTCCCACTCTTTTATAGACA 847  
 QY 772 CTTTATCTCTTAAATTTATACAGACATTTAGACAAATAGATTTATTTGGAGATGGAGA 831  
 Db 848 CTTTATCTCTTAAATTTATACAGACATTTAGACAAATAGATTTATTTGGAGATGGAGA 907  
 QY 832 CATAGTTTACAGAGATTTACTGCGAATGGAACAGCAAGAGATTACTGTTTACAG 891  
 Db 908 CATAGTTTACAGAGATTTACTGCGAATGGAACAGCAAGAGATTACTGTTTACAG 967  
 QY 892 TATGATGATCAGAAAATTTACTAGCGGCTTGGAGACACACATATCAATCTGGATAAA 951  
 Db 968 TATGATGATCAGAAAATTTACTAGCGGCTTGGAGACACACATATCAATCTGGATAAA 1027  
 QY 952 AACACATTTGGAATGCAAGCAATTTCTACAGGCAATACAGTTTCACTCTGTCTCCAG 1011  
 Db 1028 AACACATTTGGAATGCAAGCAATTTCTACAGGCAATACAGTTTCACTCTGTCTCCAG 1087  
 QY 1012 TATGATGAGAGATTTATATACAGATTCATGCAATGCAAGGTCAGAGTGGAGTGA 1071  
 Db 1088 TATGATGAGAGATTTATATACAGATTCATGCAATGCAAGGTCAGAGTGGAGTGA 1147  
 QY 1072 AATACAGGTGAATGCTTAAACAGCTTGAATTCACATTTGGAAGCAATCTGCACCTGGCT 1131  
 Db 1148 AATACAGGTGAATGCTTAAACAGCTTGAATTCACATTTGGAAGCAATCTGCACCTGGCT 1207  
 QY 1132 TTTCAATATGCGATATGCTGACCTGCTCCAAAGATGCTTCAATGCTGATAGGATATG 1191  
 Db 1208 TTTCAATATGCGATATGCTGACCTGCTCCAAAGATGCTTCAATGCTGATAGGATATG 1267  
 QY 1192 GCTCTCCCACTGACATTAACCTCGGAGGCTGCTGGGAGACACCGAGCTGCTCAAT 1251  
 Db 1268 GCTCTCCCACTGACATTAACCTCGGAGGCTGCTGGGAGACACCGAGCTGCTCAAT 1327  
 QY 1252 GTTTCAGCTTTGATGACAGATGCTTCTCATCTGGGATAGAACATTAAGGTA 1311  
 Db 1328 GTTTCAGCTTTGATGACAGATGCTTCTCATCTGGGATAGAACATTAAGGTA 1387  
 QY 1312 TGAACACAACTACTGTTGATTTGTAAGACCTTAATAGACAAACAGAGCATTCGC 1371  
 Db 1388 TGAACACAACTACTGTTGATTTGTAAGACCTTAATAGACAAACAGAGCATTCGC 1447  
 QY 1372 TGTTCAGTACAGAGGAGGCTGTAGTAGTGCATCTGACACACATATCAATTA 1431  
 Db 1448 TGTTCAGTACAGAGGAGGCTGTAGTAGTGCATCTGACACACATATCAATTA 1507  
 QY 1432 TGGGACATTAAGATGCTGCTGCTTTCAGAGTGTAGAGGCCATGAGAAATGGTCCGT 1491  
 Db 1508 TGGGACATTAAGATGCTGCTGCTTTCAGAGTGTAGAGGCCATGAGAAATGGTCCGT 1567  
 QY 1492 TGTATTCATTTGATTAACAAGAGATGCTAGTGGGGCTATATAGAAATTAAGTG 1551  
 Db 1568 TGTATTCATTTGATTAACAAGAGATGCTAGTGGGGCTATATAGAAATTAAGTG 1627  
 QY 1552 TGGGATCTTGTGCTGCTTTGGAAGGCTGCTCTGACAGGACACTGTCTACGAGCC 1611  
 Db 1628 TGGGATCTTGTGCTGCTTTGGAAGGCTGCTCTGACAGGACACTGTCTACGAGCC 1687  
 QY 1612 CTGTTCAGGATTCGGAAGAGTTTTCGACATTCAGTTTGAATTCAGATTTGTCAGT 1671  
 Db 1688 CTGTTCAGGATTCGGAAGAGTTTTCGACATTCAGTTTGAATTCAGATTTGTCAGT 1747  
 QY 1672 AGTTTCATGATGACACAACTCTCATCTGGGATCTCTTAATATGACAGTGGCCAGGT 1731  
 Db 1748 AGTTTCATGATGACACAACTCTCATCTGGGATCTCTTAATATGACAGTGGCCAGGT 1807  
 QY 1732 GAACCCCCCGTTCCTCTGACACATACCTACATCTCCAGATTAATTAACATACAC 1791

Db 1808 GAACCCCCCGTTCCTCTGACACATACCTACATCTCCAGATTAATTAACATACAC 1867  
 QY 1792 TGACCTCATCTTGGCCAGAGACCCATTAAGTTTGGGATTTAATACATTCGCCAATAC 1851  
 Db 1868 TGACCTCATCTTGGCCAGAGACCCATTAAGTTTGGGATTTAATACATTCGCCAATAC 1927  
 QY 1852 AGGATGAGCAACAACAGTAAACATCAACTACTGCCAGTTTCCCTGGACATCCAGGA 1911  
 Db 1928 AGGATGAGCAACAACAGTAAACATCAACTACTGCCAGTTTCCCTGGACATCCAGGA 1987  
 QY 1912 GCAGGCGTTTGGAGCTCTGTTGGGACACAGTTGTGTGCAATGCGGCCAGAGCTTA 1971  
 Db 1988 GCAGGCGTTTGGAGCTCTGTTGGGACACAGTTGTGTGCAATGCGGCCAGAGCTTA 2047  
 QY 1972 CTCAGCAACACTGACCTCTGAGTGTGCTATGAGAAGATGCTCTTATCAATTTGGAAT 2031  
 Db 2048 CTCAGCAACACTGACCTCTGAGTGTGCTATGAGAAGATGCTCTTATCTTTTGTGAAT 2107  
 QY 2032 GATTGGAACTTTTAAACCTCCCTCCTCTCTCTTACCTGACCTAGTTTTCCTC 2091  
 Db 2108 GATTGGAACTTTTAAACCTCCCTCCTCTCTCTTACCTGACCTAGTTTTCCTC 2167  
 QY 2092 CATTTGTTCCAGACAAAGGAGCTTAAATATATTAGTGTTCCTCCAGAAAAA 2151  
 Db 2168 CATTTGTTCCAGACAAAGGAGCTTAAATATATTAGTGTTCCTCCAGAAAAA 2227  
 RESULT 6  
 AF081887 2175 bp mRNA linear ROD 07-OCT-1999  
 LOCUS AF081887  
 DEFINITION Mus musculus ubiquitin ligase FWD1 mRNA, complete cds.  
 ACCESSION AF081887  
 VERSION AF081887.1 GI:4336326  
 KEYWORDS  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 Hatakeyama, S., Kitagawa, M., Nakayama, K., Shirane, M., Matsunoto, M., Hattori, K., Higashi, H., Nakano, H., Okumura, K., Onoe, K., Good, R.A., and Nakayama, K.-I.  
 Ubiquitin-dependent degradation of IkappaBalpha is mediated by a ubiquitin ligase Skp1/Cul1/F-box protein FWD1  
 Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3859-3863 (1999)  
 10097128  
 2 (bases 1 to 2175)  
 Hatakeyama, S. and Nakayama, K.-I.  
 Submitted (04-AUG-1998) Department of Molecular and Cellular Biology, Medical Institute of Bioregulation, Kyushu University, 3-1-1, Maidashi, Higashi-Ku, Fukuoka 812-8582, Japan  
 Location/Qualifiers  
 1..2175  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 91..1800  
 /note="similar to beta-transducin repeat-containing protein frcp, F-box protein; contains WD40 repeats"  
 /codon\_start=1  
 /product="ubiquitin ligase FWD1"  
 /protein\_id="AA017755.1"  
 /db\_xref="GI:4336327"  
 /translation="MDPAEAVLDKALKFPMNSEREDCNNGEPKRIIEKNSLRQTY NSCARCINDQETVCLSTAAKTENCVAKALANGTSMTVPKORLASAYEKEELCY KYEOWESDQVEFEHLISQCHYOHGHINSLIKMLRDEFTALPARGLHIEINI LSYDAKSAELVCEKEMRYTSDMLMKILERRVRDSDLRGLAERGGGYEYFK NKPDENAPNSFYRALYKPIIODIFIESNMGGRSH.ORIHCBSFESKYCYCLOYD DOKIVSLRNTIKIMDKSLBECKRIITJGTHVCLQYDERVITITGSSDSYRWVD NAGMFINTLIHCDAVLIHLRFNNGMYTCSKDSIAVWMASTPDIITLRYVIVGRVA VNVYDPRDKIVASGDRITKWNISTCEVFTLNHGKIGIACLOYRDLVVGSSDN TIRLMDIEGACLRVLEGHIEIVRCIRFNKRIIVSGAYDGKIKVMDLMAALDPRAPAG

BASE COUNT ORIGIN	589 a	528 c	567 g	491 t
	TLCICRTLVHSGRVRLQDFDFQIVSSHDDTILIMDFLNDPAHAAPPSPSKTYTY			
	ISR"			

Query Match	78.2%	Score 1682.2	DB 10	Length 2175
Best Local Similarly	88.5%	Pred. No. 0		
Matches 1909, Conservative	0	Mismatches 228	Indels 20	Gaps 7

Matches 1909; Conservative 0; Mismatches 228; Indels 20; Gaps 7;

Matches 1909; Conservative 0; Mismatches 228; Indels 20; Gaps 7;

Oy	1	TGCGTTGGCTGGGCGCTGGACACAAAGGCGCGCCGCGGAGAGCGAGCCACCTGGCC	60
Db	22	TGCGTTGGCTGGGCGCTGGACAAAGGCGCGCCGCGGAGAGCGAGCCACCTAGTGC	81
Oy	61	TGCGCGATTATGAGACCCGGCGAGGCGGTGCTGCAAGAAAGGCACTCAAGTTTATGAT	120
Db	82	CGGGGAGATTATGAGACCCGGCGAGGCGGTGCTGCAAGAAAGGCGCTTAAAGTTATGAT	141
Oy	121	TCCTCAGAGAGAGAGACTGTATATATGGCGAAACCCCTAGTAAATATATACCAGAGAG	180
Db	142	TCCTCAGAGAGAGAGAGACTGTATATATGGCGAAACCCCTAGTAAATATATACCAGAGAG	201
Oy	181	AATTCACCTTAGACAGACATACAAACAGCTGTGCGACACTCTGCTTAAACCAAGAACGTA	240
Db	202	AATTCACCTTAGACAGACTTACAAACAGCTGTGCGACACTCTGCTTAAACCAAGAACGTA	261
Oy	241	TGTTTAGCAAGAGACTGCTATGAAAGCTGAGATTTGTTGGCCAAAACAACCTTGCCAT	300
Db	262	TGTTTAGCAAGAGACTGCTATGAAAGCTGAGATTTGTTGGCCAAAACAACCTTGCCAT	321
Oy	301	GGCAGCTCCAGTATGATTTGTGCCAAGCAAGGAAACTCTCAGCAAGCTATGAAAGGAA	360
Db	322	GGCAGCTCCAGATGATTTGTGCCAAGCAAGGAAACTCTCAGCAAGCTATGAAAGGAA	381
Oy	361	AAGGAACTGTGTCAATACCTTTGGAGAGTGGTGCAGTGCAGATCAGATCAATGGAAATTGG	420
Db	382	AAGGAACTGTGTCAATACCTTTGGAGAGTGGTGCAGTGTGATGATCAATGGAAATTGGTA	441
Oy	421	GAACATCTTATATCCCAAATGTGTCTTACCAACATGGGACATAACTCGATCTTAA	480
Db	442	GAACACCTTATATCCCAAATGTGTCTACACAGCATGGGACATCACTCTACTTAA	501
Oy	481	CCATATGTGCAAGAAATTTCAATAACTGCTGCCAGCGTGGGGAATGGATCATATCGCT	540
Db	502	CCATATGTGCAAGGGAATTTCAATAACTGCACTGCCAGCGTGGGGAATGGATCATATCGCT	561
Oy	541	GAGAACTTCTGTCATACCTGTGATGCCAAATCACTATGTGCTGCTGCACTTGTGTCAAG	600
Db	562	GAAACATTTCTGTCACTACTTGGAGGCGCAAGTCACTGTGTGCTGCTGAGCTCGTGTCAAG	621
Oy	601	GAATGTGATCCGAGTGCATCTGTATGCGATGTCTGTGGAAGAAAGTTATCGAGAAATGTC	660
Db	622	GAATGTGATCCGCTGTGAGAGGCGCTGTGCGAAGCAAGAAAGTATGAGAAAGTATGTC	681
Oy	661	AGGACAGATTTCTGTGAGAGGCGCTGTGCGAAGCAAGAAAGTATGAGAAAGTATGTC	720
Db	682	AGGACAGACTCTGTGTGCGAGGCGCTGTGCGAAGCAAGAAAGTATGAGAAAGTATGTC	741
Oy	721	AAAAACAACCTCTCTGACGCGGAATGCTCTCCCACTCTTTTATAGAGCATTTATCT	780
Db	742	AAAAACAACCTCTCTGATGAAAGCGCTCTCCCACTCTTTTATAGAGCGCTTATCT	801
Oy	781	AAATTTATACAGCATTTGAGCAATATGATATGATTTGAGATGTGGACACATAGTTTA	840
Db	802	AAATTTATACAGCATTTGAGCAATATGATTTGAGATGTGGACACATAGTTTA	861
Oy	841	CAGAGAACTTCACTGCCGAGTGAAGAAAGCAAGCAAGGACTTACTGTTTACAGTATGATGAT	900
Db	862	CAGAGAACTTCACTGCCGAGTGAAGAAAGCAAGGAGGCTTACTGTTTACAGTATGATGAT	921
Oy	901	CAGAAATTTACTAAGCGGCTTTGAGACACACATCAAGATCTGGGATTAAGAAACACTTG	960
Db	922	CAGAAATTTACTAAGCGGCTTTGAGACACACATCAAGATCTGGGATTAAGAAACACTTG	981

Qy	961	GAATGCAAGGAAATTTCTCAGACGCCATACAGSTTCAAGTCCCTGCTCCACAGTAAATGATGAG	1020
Db	982	GAATGCAAGGGAGATTCTCAGGGCCACACAGGGCTCCGTCTGTCTGCGACAGTACGATGAG	1041
Qy	1021	AGAGTGAATCAATACAGAGATCATCGATTCTCCACAGGTCACAGTGTGGATGTAAATACAGT	1080
Db	1042	AGGGTGAATCATCAGAGGCTCTCTCAGACTCCACCCTCACAAGTGTGGATGTAAATACAGG	1101
Qy	1081	GAATGCTAAACACGTTTGATTCACCATTTGTGAAGCAGTTCTTGACACTTGCCTTTCATATAT	1140
Db	1102	GAGATGCTAAACACATTTGATTCACCATTTGTGAAGCAGTTCTTGACACTTGCCTTTCATATAT	1161
Qy	1141	GGCATGATGGGACCTGCTCCAAAGATTCCTTCATTTGCTGATATGAGGATATAGGCTCCCA	1200
Db	1162	GGCATGATGGGACCTGCTCCAAAGATTCCTTCATTTGCTGATATGAGGATATAGGCTCCCA	1221
Qy	1201	ACTGCATTTACCCCTCCGAGGAGTGTGTCGAGACACGAGCTGCTTCATATGTTGTAGAC	1260
Db	1222	ACTGCATTTACCCCTCCGAGGAGTGTGTCGAGACACGAGCTGCTTCATATGTTGTAGAC	1281
Qy	1261	TTTGTATGACAAGTACATTTGTTCTGATCTGGGATATAGACTATTAAGATATGGAACACA	1320
Db	1282	TTTGTATGACAAGTACATTTGTTCTGATCTGGGATATAGACTATTAAGATATGGAACACA	1341
Qy	1321	AGTATTTGTGAATTTGTAAAGACCTTAAATGAGACAAACAGGAGTTCGCTGTTTGACG	1380
Db	1342	AGTATTTGTGAATTTGTAAAGACCTTAAATGAGACAAACAGGAGTTCGCTGTTTGACG	1401
Qy	1381	TACAGGACAGGCTGATAGTGAAGTGCCTATCTGACACACTATTCAGATTAAGGACATA	1440
Db	1402	TACAGGACAGGCTGATAGTGAAGTGCCTATCTGACACACTATTCAGATTAAGGACATA	1461
Qy	1441	GAATGTGTGATGTTACAGAGTCTTAAGAGCCATAGAGAGATTTGGTGGCTGTAAATGCA	1500
Db	1462	GAATGTGTGATGTTACAGAGTCTTAAGAGCCATAGAGAGATTTGGTGGCTGTAAATGCA	1521
Qy	1501	TTTGTATTAACAAGAGATAGTACAGTGGGCTATGATGAGAAAAATTAAGTGTGGATCTT	1560
Db	1522	TTTGTATTAACAAGAGATAGTACAGTGGGCTATGATGAGAAAAATTAAGTGTGGATCTT	1581
Qy	1561	GTGGCTGCTTTTGACCCCCCGTGTCTCTGACAGGACACTGTCTACAGACCCCTGTGGAG	1620
Db	1582	ATGGCTGCTTTTGACCCCCCGTGTCTCTGACAGGACACTGTCTACAGACCCCTGTGGAG	1641
Qy	1621	CATTCGGAAGATTTTTCGACTACATTTGATGAATTCACAATTTGTGACAGTTCACAT	1680
Db	1642	CATTCGGAAGATTTTTCGACTACATTTGATGAATTCACAATTTGTGACAGTTCACAT	1701
Qy	1681	GATGACACATCTCATCTGSGACTTGCCTAAATGATCAGCTGAGCCCAAGCTGAACCCCC	1740
Db	1702	GATGACACATCTCATCTGSGACTTGCCTAAATGATCAGCTGAGCCCAAGCTGAACCCCC	1761
Qy	1741	CGTCCCTCTTCGACAACTTACCTACATCTCCAGATTAATAACATTAACATTAACATTAAT	1800
Db	1762	CGTCCCTCTTCGACAACTTACCTACATCTCCAGATTAATAACATTAACATTAACATTAAT	1821
Qy	1801	ACTTGCCCGACCCATTAAAGTTGCGGTATTTAAAGTATTCGCCAATACAGATGAGC	1860
Db	1822	AATTGCCCGACAGTTCGTTAAAGTTGCGGTATTTAAAGTATTCGCCAATACAGATGAGC	1881
Qy	1861	AACACAGTAAACATCAAACTACTGCCCAGTTTCCCTG--GACTAGCCGAGAGAGAGGCG	1918
Db	1882	-----AACACAGTAAACATCAAACTACTGCCCAGTTTCCCTG--GACTAGCCGAGAGAGAGGCG	1932
Qy	1919	TTTGTAGACCTGTTGGGACACAGTTGGTCTGAGTGGGCGGAGCGGTCTACTACACA	1978
Db	1933	TTTGTAGACCTGTTGGGACACAGTGGTCTGAGTGGGCGGAGCGGTCTACTACACA	1990
Qy	1979	CAACTACTGCTTCACTAGTCTGCTATCAGAAAGATGTCTTCTATCAATTTGTAATGATGGA	2038
Db	1991	--ACCGGCTGCGCTAGTGTGCTATCAGAAAGATGTCTT--TATCTTGTGTAATGATGGA	2047



QY	1372	TGTTTCAGACGACGGGACAGAGCGGTGAGTAGTAGGGGTCATCTGACACACATATACAGATTA	1431
Db	1411	TGTTTCAGACGACGGGACAGAGCGGTGAGTAGTAGGGGTCATCTGACACACATATACAGATTA	1470
QY	1432	TGGGACATACAAAGTGGTGCATGTTTACGAGTGTAGAAAGGCCATGAGAAATTTGGTGGT	1491
Db	1471	TGGGACATACAAATGTGGTGCATGTTTACGAGTGTAGAAAGGCCATGAGAAATTTGGTGGT	1530
QY	1492	TGTAATTCGATTTGATATACACAGAGATAGTCAGTGGGCGCCATCATGAGAAAAATTAAGTG	1551
Db	1531	TGTAATTCGATTTGATATACACAGAGATAGTCAGTGGGCGCCATCATGAGAAAAATTAAGTG	1590
QY	1552	TGGGATCTTGTGGCTGCTTTGGAACCCCGCTCTCTCAGAGGACACATCTGTACGGAGC	1611
Db	1591	TGGGATCTTGTGGCTGCTTTGGAACCCCGCTCTCTCAGAGGACACATCTGTGTACGGAGC	1650
QY	1612	CTTGTGAGCATTTCCGGAGAGAGTTTTCAGACTACAGTTTATGATTCGAGATGTGAGT	1671
Db	1651	CTTGTGAGCATTTCCGGAGAGAGTTTTCAGACTACAGTTTATGATTCGAGATGTGAGT	1710
QY	1672	AGTTCACATGATGACACAAATCCATCAGTGGGACTTCCATAATGATCCAGCTGCCAAGCT	1731
Db	1711	AGTTCACATGATGACACAAATCCATCAGTGGGACTTCCATAATGATCCAGCTGCCAAGCT	1770
QY	1732	GAACCCCCCGTTCCCTTCTCGAACATACACCTACATCTTCCAAATTA	1779
Db	1771	GAACCCCCCGTTCCCTTCTCGAACATACACCTACATCTTCCAAATTA	1818
RESULT 8			
LOCUS	BC003989	2081 bp	mRNA linear ROD 07-AUG-2002
DEFINITION	Mus musculus, beta-transducin repeat containing protein, clone		
ACCESSION	MG:7517 IMAGE:3491843, mRNA, complete cds.		
VERSION	BC003989		
KEYWORDS	BC003989.1 GI:13278339		
SOURCE	MGC.		
ORGANISM	house mouse.		
	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 2081)		
AUTHORS	Strausberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC) Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: <a href="http://mgc.ncl.nih.gov">http://mgc.ncl.nih.gov</a>		
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Sequencing group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: <a href="http://www.shgc.stanford.edu">http://www.shgc.stanford.edu</a> Contact: (Dickson, Mark) <a href="mailto:mdcpaxil@stanford.edu">mdcpaxil@stanford.edu</a> Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.		
FEATURES			
source	1. 2081 /organism="Mus musculus" /db_xref="taxon:10090" /map="G2BCH 11" /clone="MGC:7517 IMAGE:3491843"		

Query Match	Best Local Similarity	88.5%;	Pred. No. 0;	Matches 1853;	Conservative	0;	Mismatches 221;	Indels 20;	Gaps 7;
64	CGGATTATGACCCGGCCGAGCGGGTGCCTGCGCAAGAGAAAGGCATCAAGTTATGAAATGCC	123							
1	CGGATTATGACCCGGCCGAGCGGGTGCCTGCGCAAGAGAAAGGCATCAAGTTATGAAATGCC	60							
124	TCGAGAGAGAGACCTGTATATATGCGCAACCCCTTGAAGATATATACAGAGAAAGAT	183							
61	TCGAGAGAGAGAAAGCTGTATATATGCGCAACCCCTTGAAGATATATACAGAGAAAGAT	120							
184	TCAGTTAGACAGACATACAAACAGCTGTGCGCAGACCTTGCTTAACCAAGAAACAGTATGT	243							
121	TCAGTTAGACAGACATACAAACAGCTGTGCGCAGACCTTGCTTAACCAAGAAACAGTATGT	180							
244	TTAGCAAGACAGCTGTATGAAAGCTGGAATTTGTGTGGCCAAAACAACATTTGCCAATGGC	303							
181	CTBACAGACAGCTGTATGAAAGCTGGAATTTGTGTGGCCAAAACAACATTTGCCAATGGC	240							
304	ACTTCAGATGATGTGTGCCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAGAAAG	363							
241	ACTTCAGATGATGTGTGCCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAGAAAG	300							
364	GAACCTGTGTCTCAAACTTTTGGAGAGTGGTTCAGAGCTCAGATCAAGTGAATTTGTGGAA	423							
301	GAGCTGTGTCTCAAACTTTTGGAGAGTGGTTCAGAGCTCAGATCAAGTGAATTTGTGGAA	360							
424	CATCTTATATCCCAAAATGTGTCTATTACCAACATGGGCATTAACCTGTATCTTAAACCT	483							
361	CACCTTATATCCCAAAATGTGTCTATTACCAACATGGGCATTAACCTGTATCTTAAACCT	420							
484	ATGTTGCAAGAGATTTTCATTAACCTGCTCTGCGCCAGAGCTGGGAGATTTGATTCATTCGCGAG	543							
421	ATGTTGCAAGAGATTTTCATTAACCTGCTCTGCGCCAGAGCTGGGAGATTTGATTCATTCGCGAG	480							
544	AACATTTCTGCATACCTGATGCGCAAAATCACTATGTCTCTCTGAACTTTGTGTGCAAGAA	603							
481	AACATTTCTGCATACCTGATGCGCAAAATCACTATGTCTCTCTGAACTTTGTGTGCAAGAA	540							
604	TGTTACCGAGTGAACCTGTGATGGCATCTGTGTGGAAGAAGCTTATCCGAGAAATGGTAGG	663							
541	TGTTACCGAGTGAACCTGTGATGGCATCTGTGTGGAAGAAGCTTATCCGAGAAATGGTAGG	600							
664	ACGAGATCTCTGTGAGAGAGCCTGTGCAAGCAAGAGAGATGGGAGCAAGTATTTATTCAAA	723							
601	ACGAGATCTCTGTGAGAGAGCCTGTGCAAGCAAGAGAGATGGGAGCAAGTATTTATTCAAA	660							



QY 130 AGAGAGACTGTAATTAATGAGGAAACCCCTAGGAATATACAGAGAAATTCATT 189  
 DB 61 AGAGAGAGCTGTAATTAATGAGGAAACCCCTAGGAATATATACAGAGAAATTCATT 120  
 QY 190 AGACAGACTATACAGAGCTGTGCGAGACTGTCTTAAACAGAAACAGATGTTAGCA 249  
 DB 121 AGACAGACTATACAGAGCTGTGCGAGACTGTCTTAAACAGAAACAGATGTTAGCA 180  
 QY 250 AGCAGCTGTATGAGAACTGAGAAATTTGTGCGCAAAACAACTGCCAATGGCACTTC 309  
 DB 181 AGCAGCTGTATGAGAACTGAGAAATTTGTGCGCAAAACAACTGCCAATGGCACTTC 240  
 QY 310 AGTATGATTTGTCGCCAGAGACGAAACTCTGACAGACTATGAAAGAAAGAAAGAC 369  
 DB 241 AGCATGATTTGTCGCCAGAGAGGAAACTCTGACAGACTATGAAAGAAAGAAAGAC 300  
 QY 370 TGTGTCAATTAATTTGAGAGAGTGTGCGAGATCAATCAAGTGAATTTGTGAGACATCT 429  
 DB 301 TGTGTCAATTAATTTGAGAGAGTGTGCGAGATCAATCAAGTGAATTTGTGAGACATCT 360  
 QY 430 ATATCCCAATGTCATTAATCAACATGCGACATTAACCTGTATCTTAAACCTATGTTG 489  
 DB 361 ATATCCCAATGTCATTAATCAACATGCGACATTAACCTGTATCTTAAACCTATGTTG 420  
 QY 490 CAGAGAGATTTGATTAATGCTGTGCGAGCTCGGGATTTGATCAATGCTGAGAACT 549  
 DB 421 CAGAGAGATTTGATTAATGCTGTGCGAGCTCGGGATTTGATCAATGCTGAGAACT 480  
 QY 550 CTGTCTACCTGTGAGAGCCAAATCACTATGCTGCTGAACCTGTGTGCGAGAAATGTTG 609  
 DB 481 CTGTCTACCTGTGAGAGCCAAATCACTATGCTGCTGAACCTGTGTGCGAGAAATGTTG 540  
 QY 610 CGAGTACCTGTGAGAGCCAAATCACTATGCTGCTGAACCTGTGTGCGAGAAATGTTG 669  
 DB 541 CGAGTACCTGTGAGAGCCAAATCACTATGCTGCTGAACCTGTGTGCGAGAAATGTTG 600  
 QY 670 TCTCTGTGAGAGCCCTGTGCGAGAAAGAGATGCGGACAGTATTATTCAAAAACAA 729  
 DB 601 TCTCTGTGAGAGCCCTGTGCGAGAAAGAGATGCGGACAGTATTATTCAAAAACAA 660  
 QY 730 CTTCTGTGAGAGCCCTGTGCGAGAAAGAGATGCGGACAGTATTATTCAAAAACAA 789  
 DB 661 CTTCTGTGAGAGCCCTGTGCGAGAAAGAGATGCGGACAGTATTATTCAAAAACAA 720  
 QY 790 CAGAGATTTGAGAAATGATGAGATGAGAGATGAGAGATGATTTACAGAGAT 849  
 DB 721 CAGAGATTTGAGAAATGATGAGATGAGAGATGAGAGATGATTTACAGAGAT 780  
 QY 850 CACTGCCGAGAGAAACAGAAAGAGATTTACTGTTTACAGATATGATCAGAAATA 909  
 DB 781 CACTGCCGAGAGAAACAGAAAGAGATTTACTGTTTACAGATATGATCAGAAATA 840  
 QY 910 GTAGAGGCTGTGAGAGCAACAGATGATGAGATGAGATGAGATGAGATGAGATGAG 969  
 DB 841 GTAGAGGCTGTGAGAGCAACAGATGATGAGATGAGATGAGATGAGATGAGATGAG 900  
 QY 970 CGAATTTCTACAGGCGCATACAGGTTGAGTCTGCTCTGCTCAGATGATGAGAGATGAT 1029  
 DB 901 CGAATTTCTACAGGCGCGCATACAGGTTGAGTCTGCTCTGCTCAGATGATGAGAGATGAT 960  
 QY 1030 ATACAGAGATGATGAGATTTGAGAGTGTGAGATGAGATGAGATGAGATGAGATGAG 1089  
 DB 961 ATACAGAGATGATGAGATTTGAGAGTGTGAGATGAGATGAGATGAGATGAGATGAG 1020  
 QY 1090 AACAGGTTGATTTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCT 1149  
 DB 1021 GACAGATTTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAG 1080  
 QY 1150 GTGACCTGTCTCAAAAGATGTTGATGAGAGTGTGATGAGAGTGTGATGAGAGTGTG 1209  
 DB 1081 GTGACCTGTCTCAAAAGATGTTGATGAGAGTGTGATGAGAGTGTGATGAGAGTGTG 1140

QY 1210 ACCCTCCGAGAGGCTGCTGTGCGAGACCCGAGCTGCTGCTCAATGTTGTAGACTTTGATGAC 1269  
 DB 1141 ACCCTCCGAGAGGCTGCTGTGCGAGACCCGAGCTGCTGCTCAATGTTGTAGACTTTGATGAC 1200  
 QY 1270 AAGTACATTTGTTCTGATCTGTGAGATGAGAACTATTAAGATGATGAACACAACTACTTGT 1329  
 DB 1201 AAGTACATTTGTTCTGATCTGTGAGATGAGAACTATTAAGATGATGAACACAACTACTTGT 1260  
 QY 1330 GAATTTGTAAGGACCTTAAATGAGACAAACGAGGATGCTGTTTTCAGTACAGAGAC 1389  
 DB 1261 GAATTTGTAAGGACCTTAAATGAGACAAACGAGGATGCTGTTTTCAGTACAGAGAC 1320  
 QY 1390 AGCCTGTGATGAGTGTGCTATCTGACAAACATCTGATTTATGAGATGAGATGAGTGT 1449  
 DB 1321 AGCCTGTGATGAGTGTGCTATCTGACAAACATCTGATTTATGAGATGAGATGAGTGT 1380  
 QY 1450 GCATGTTTACAGAGTGTGTAAGAGGCTATGAGAAATGAGTGTGCTGATTTGATGATGAT 1509  
 DB 1381 GCATGTTTACAGAGTGTGTAAGAGGCTATGAGAAATGAGTGTGCTGATTTGATGATGAT 1440  
 QY 1510 AAGAGATGATGAGTGTGAGGCTATGATGAGAAATTAAGTGTGAGTGTGCTGCT 1569  
 DB 1441 AAGAGATGATGAGTGTGAGGCTATGATGAGAAATTAAGTGTGAGTGTGCTGCTGCT 1500  
 QY 1570 TTGAGACCCGCTGCTCTGCGAGAGACACTGCTGCTAGGACCCCTGAGGACATTCGGA 1629  
 DB 1501 TTGAGACCCGCTGCTCTGCGAGAGACACTGCTGCTAGGACCCCTGAGGACATTCGGA 1560  
 QY 1630 AGAGTTTTCAGTACAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1689  
 DB 1561 AGAGTTTTCAGTACAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
 QY 1690 ATCTCTATCTGAGAGCTTCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1749  
 DB 1621 ATCTCTATCTGAGAGCTTCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
 QY 1750 TCTGACATCACTACATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1809  
 DB 1681 TCTGACATCACTACATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
 QY 1810 GGACCCATTAAAGTTGCGGATTTAAAGTATGATGATGATGATGATGATGATGATGATGAT 1869  
 DB 1741 GGACCCATTAAAGTTGCGGATTTAAAGTATGATGATGATGATGATGATGATGATGATGAT 1791  
 QY 1870 AACCAATCAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1927  
 DB 1792 AACCAATCAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1851  
 QY 1928 CCTGTTGGAGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1987  
 DB 1852 CCTGTTGGAGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1907  
 QY 1988 GCTTCAGTGTCTATACAGAGATGCTTCTATCAATTTGATGATGATGATGATGATGAT 2047  
 DB 1908 GCTTCAGTGTCTATACAGAGATGCTTCTATCAATTTGATGATGATGATGATGATGAT 1966  
 QY 2048 CTTCCCTCTCC 2057  
 DB 1967 CTTCCCTCTCC 1976

RESULT 10  
 AF099932 1712 bp mRNA Linear ROD 21-JAN-1999  
 LOCUS AF099932 Mus musculus beta-Tcp protein E3RS-Ikappab mRNA, complete cds.  
 DEFINITION AF099932  
 ACCESSION AF099932.1 GI:4008019  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus.  
 Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1 (bases 1 to 1712)

AUTHORS	Varon, A., Hatzubai, A., Davis, M., Lavon, I., Amit, S., Manning, A.M., Andersen, J.S., Mann, M., Mercurio, F. and Ben-Neriah, Y.	
TITLE	Identification of the receptor component of the IkappaBalpha-ubiquitin ligase	
JOURNAL	Nature 396 (6711), 590-594 (1998)	
MEDLINE	99075339	
PUBMED	9859996	
REFERENCE	2 (bases 1 to 1712)	
AUTHORS	Varon, A., Hatzubai, A., Mercurio, F., Manning, A.M., Andersen, J.S., Mann, M. and Ben-Neriah, Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (19-OCT-1998) Immunology, Hebrew University of Jerusalem, Ein Karem, Jerusalem 91120, Israel	
FEATURES	Location/Qualifiers 1..1712 1..1712 "organism="Mus musculus" /db_xref="taxon:10090" 1..1710 /note="beta-transducin repeat-containing protein; receptor component of IkBa-ubiquitin ligase" /codon_start=1 /product="beta-Trip protein E3NS-IkappaB" /protein_id="AABD08701.1" /db_xref="GI:4008020" /translation="MDPAEALVDEKALKFNKSSEREDCNNGEPPRKIIPEKNSIRQTYNSCAALINQETVLCITSTAMKTEVQENCAKLANGSSATVFKRKLSSYKEKELCYKFEWSSDOVEFEVHELISQMHQHOHINSYKLPMDRPFITALPRGIDHIAENILSYDANSLCAAEIVCEKEMWYRHTSDGMKMLIEMWYETDLSLMLGARERMGQYLKPKPDENAPNPSFVALPKIIDIETIESHWGGRHSILOIHRSEPKVYLTQTDOKTISGLRDMTIFKIMWSTLECKRIILTHGHSVLCIYQDERVITITGSSDSVWVWVYVVDPEDKYIVSASGDKTIFKVMNISTCEFRVTLNGHKRIACIQIDRLRLVSSSDNITRIMDIPCGACLRVLEGHEELVRCIRPDNKRIVSAYDGKIKYWDLMALDPRAPAGTLCITLVEHSGRWFRILQPDFEFOIVSSSHDPTILIMFDLPNAAHEPPRSPRYTYISR"	
BASE COUNT	469 a 399 c 453 g 391 t	
ORIGIN		
Query Match	67.5%; Score 1451.2; DB 10; Length 1712;	
Best Local Similarity	90.5%; Pred. No. 0;	
Matches 1549: Conservative	0; Mismatches 163; Indels 0; Gaps 0;	
70	ATGAGCCGGCCGAGGCGGTGCTGCAAGAGGACATCAAGTTATATTCCTCAGAG	129
Db	1 ATGGACCGGCGCGAGCGGTGCTGCAAGAGAAAGCGCTTAAGTTATATTCCTCAGAG	60
QY	130 AGAGAGACTGTAATATATGCGGAACCCCTTGAAGAATATATCCAGAGAAATTCATT	189
Db	61 AGAGAGACTGTAATATATGCGGAACCCCTTGAAGAATATATCCAGAGAAATTCATT	120
QY	190 AGACAGACATCAACAGCTGTGCCAGACTCTGCTTAAACCAAGAACGVTGTTAGCA	249
Db	121 AGACAGACTTCAACAGCTGTGCCAGCGTTTGCAATAAACCAAGAGACAGTATGCTTACA	180
QY	250 AGCACTGCTATGAAAGCTGAGAAATTTGTGTGCCAAACAAACCTTGGCAATGGCACTTCC	309
Db	191 AGCACTGCTATGAAAGCTGAGAAATTTGTGTGCCAAAGCAAACTTGGCAATGGCACTTCC	240
QY	310 AGTATGATTGTGCCCAACACAGGAAACTCTCAGACAGTATGAAAAGAAAAGAACTG	369
Db	241 AGCATGATTGTGCCCAACAGGAAACTCTCAGACAGTATGAAAAGAAAAGAACTG	300
QY	370 TGTGTCAAAATCTTTGAGCACTGGTCCAGAGTCAAGTGAAGTAATTTGTGGAACATCTT	429
Db	301 TGTGTCAAGTATTTTGGACAGTGGTCCAGAGTCTGATCAAGTGAATTTTGTGGAACATCTT	360
QY	430 ATATCCCAAAATGTGCAATTTACCAACATATGGGACATATAACTGATATCTTAAACCTATGTTG	489
Db	361 ATATCCCAAAATGTGCTACTACCAACATATGGGACATATCAATCTTAAACCTATGTTG	420
QY	490 CAGAGAGATTTCAATACCTGCTCTGCCAGCTCGGGGATTTGGATCATATACGCTGGAAACATT	549
Db	421 CAGAGGATTTCAATACCTGCACTGCCAGACAGGGGTGTGGACCAACATCGCTGGAAACATT	480

QY	550	CTGTCACTACCTGGATGCGAAATACATATGTGCTGCTGAACCTTGTGTGCAAGGAATGGTAC	609
Db	481	CTGTCACTACTTTGGACGCGCAAGTACATCTGTGTGCTGTGAGCTGTGTGCAAGGAATGGTAC	540
QY	610	CGAGTGAACCTCTATGATGGCATGTGCTGTGGAAAGAGCTTAATCAAGAGAAATGCTAGAGAT	669
Db	541	CGGGTGAACGTGCGAGCGGCACTGTGTGGAAAAAGCTCATCGAGAGAGATGCTAGAGACGAC	600
QY	670	TCTCTGTGGAGAGGAGCCTGGCAGAACGAGAGATGGGGACAGTATTTATTCAAAAACAAA	729
Db	601	TCTCTGTGGAGAGAGCCTGGCAGAGCGAGAGAGGCTGGGGACAGTACTTAATTCAAAAACAAA	660
QY	730	CTCTCTGACGGAGATCTCTCTCCCAACTTTTATTAAGACACTTTATTCCTAAAAATATA	789
Db	661	CTCTCTGATGAGAAAGCCTCTCTCCAACTCCTTTATTAAGAGCGCTTTATTCCTAAAAATATA	720
QY	790	CAAGACATTGAGACATATAGATCTAATTTGAGATGTGGAAACATAGTTTACAGAGATT	849
Db	721	CAAGACATTGAGACATATAGTCCATTGTGGAGATGTGGGCGACATAGTTTACAGAGATT	780
QY	850	CATGTGCCGAAGTGAACCAAGCAAGAGATTTACTGTTTACAGTATGATGATCAAGAAATA	909
Db	781	CATGTGCCGAGTGAACCAAGTAAAGGGTTTACTGTTTACAGTATGAGCAGCAAGAAATA	840
QY	910	GTAAGCGGCTTTCGAGACACACAAATCAAGATCTGGGTAAAAACAATTTGGAATGCAAG	969
Db	841	GTCAGCGGCTTTCGAGACACACCATCAAGATCTGGGTAAAAAGCACACGTGAATCCAG	900
QY	970	CGAATCTTCACAGGCGCATACAGGTTCACTCCTCTGTCTCCAGTATGATGAGAGATGATC	1021
Db	901	CGGATCTTCACGGGCGCACAGGGGCTCGTCTGTGTGTGAGTACGATGAGAGAGGATGATC	960
QY	1030	ATPACAGATATCCGATTCGATTCACGGTACAGCTGTGGGATTAATAACAGGTGAAATGCTA	1081
Db	961	ATPACAGGCTCTCAGACTCCACCGTACAGATGTGGGATTAATATCAGGTGAGTACTTA	1021
QY	1090	AACAGGTGATCACCATTTGTGAAGCAGTTTGTGACTTGGCTTCAATAATGAGCATATG	1141
Db	1021	AACACATTTGATTCACCATGTGAAGCGGTTCTGTGCACTTGGCTTCAATAATGAGCATATG	1080
QY	1150	GTGACCTGTCTCAAAAGATCGTTCATTTGCTGTATGGGATATGGGCGTCCCAACTGACAT	1201
Db	1081	GTGACCTGTCTCAAAAGACGTTTCATGTGCTGTGTGGGATATGGCTTCCCAACTGATCATC	1141
QY	1210	ACCCCTCCGAGGGGTGTGTGTGGACACCGCAGCTGTCTCAATGTTGTAGACTTTGATGAC	1261
Db	1141	ACCCCTCAGAGGGGTGTGTGTGGACACCGCAGCTGTCTCAATGTTGTAGACTTTGATGAC	1200
QY	1270	AGTACATTTGTTTCGATCTGGGGGATGAGAACTATAAGGATGGAACACACAGTACTGT	1321
Db	1201	AGTACATCTGTTTCTGCTCTCTGTGGAGATGAGAACATTAAGGTGTGGAAACCAAGTACTGT	1260
QY	1330	GAATTTGTAAGGACCTTAAATGAGACACAAACGAGGACTTGCTGTTTGAGATGACGAGAC	1389
Db	1261	GAATTCGTAAGGACCTTAAATGAGCACAAACGATGGCATGGCGTGTGGAGTACAGAGAC	1320
QY	1390	AGGCTGGATGAGATGGGCTCATCTGACAAACACTATACATTAATGGACATACAGATGTGT	1449
Db	1321	AGGCTGTGTGGAGGGCTCTCTCTGACAAACCATATAGGCTGTGGACACATTAAGGTGTGGA	1380
QY	1450	GCATGTTTACAGATGTTAAGAGGCGCATGAGGAATGTGTGCTGTGATTCGATTTGATATAC	1509
Db	1381	GCATGCTTTCGAGATGTGGAGGGCGCATGAGAGATGTGTACCTGCATTCGATTTGATATAC	1440
QY	1510	AAGAGATAGTCAGTGGGGCTTATGATGAAAAATTAAAGTGTGGGATCTTGTGGCTCT	1569
Db	1441	AAAAGGATAGTACCGGAGGCTTATGATGGAAAAATTAAAGTGTGGGATCTTATGGCTCT	1500
QY	1570	TTGAGACCCCGTGTCTCTGTGAGGAGACACTGTGTAACGACACCTTGTGGAGCATTTCCGGA	1629
Db	1501	TTGACCCCGGTGTCTCTGTGAGAGGACTCTCTGTCTGTGGACACTTGTGTGAGCATTTTGA	1560

Qy	1630	AGAGTTTTCCACATACAGATTTGATCAATTCAGATGTCAGAGTTCACATGATGACACA	1689
Db	1561	AGAGTTTTCCACCTCCACATTTGATGAATTCAGATTTGACAGAGTTCACATGATGACACA	1620
Qy	1690	ATCCCATCTGGAGATTCCTAATATGATCCAGCTGCGCAAGCCGCCCGCTGCCCT	1749
Db	1621	ATTCCTCATCTGGAGATTCCTGATGATATCCAGCTGCTCAGCTGGAACCGCCCGCTGCCCT	1680
Qy	1750	TCTCGACATACACCTACATCTCCAGATAAAT	1781
Db	1681	TCTGGACATACACTACATCTCCAGATAAAT	1712

RESULT 11	
AF112979	1710 bp mRNA linear ROD 02-MAR-1999
LOCUS	
DEFINITION	Mus musculus beta-transducin repeat containing protein mRNA,
	complete cds.
ACCESSION	AF112979
VERSION	AF112979.1 GI:4140717
KEYWORDS	
SOURCE	Mus musculus.
ORGANISM	Mus musculus.

REFERENCE AUTHORS TITLE	JOURNAL MEDLINE PUBMED
1 (bases 1 to 1710) Spencer, E., Jiang, J. and Chen, Z.-J. Signal-induced ubiquitination of IkappaBalpha by the F-box protein Slimb/Beta-TTCP	98145465 9990853
2 (bases 1 to 1710) Chen, Z.-J. Direct Submission Submitted (10-DEC-1998) Molecular Biology and Oncology, UT	

FEATURES	Location/Qualifiers
source	1. .1710

**CDS**

/function=ubiquitinates phosphorylated Ikba"  
/note=beta-Trcp; F-box protein; Ikb-ubiquitin ligase;  
substrate recognition subunit of SCF complex; similar to  
Homo sapiens beta-Trcp and Drosophila melanogaster Slimb"

```
/codon_start=1  
/product="beta-transducin repeat containing protein"  
/protein_id="AAD04181.1"  
/db_xref="GI:4140718"  
/translation="MDKADGKQVGVYRSGEEDDQNNNEDEDDRTTPEKNSDOT"
```

/LEADSLALON- MPEAFALUQALANP MINSSELEKANOEF FENL I PUNKEL  
 NSCARLITNOEVLCTSTANTKENCAKAKANGSSIMYIPKORLSASVEKEJLQ  
 KTFEOMESDOYEVEFELILSOMCHYOHGHINSY LKPMLODFITLAPARGLDHAEN  
 LISDASISALAEVCEWYRTDSDMLMKKLIERVARTDLSLNGIAERAGGOLV  
 WYRDDENADNNEEYVALYKRTITSDMLSNMWRGHSVORTHSGSEFEGGVALOY

[illegible]

ISR<sup>a</sup> 399 C 453 G 389 T

67.18;	Score 1442.8;	DB 10;	Length 1710;
ity 90.28;	Pred. No. 0;		
Derivative	0: Mismatches 167;	Indels 0;	Gaps 0

QY 70 ATGACACCCGGCCGAGGGCGTCTCTCAAGAGAGAACGCACTCAAGTTTATGAATTCCTCAGAG 129  
1 ATGACACCCGGCCGAGGGCGTCTCTCAAGAGAGAACGCACTCAAGTTTATGAATTCCTCAGAG 60  
Db

QY 130 AGAGAGACGCTAATAATGGCCACCCTTAGSAGATATAACCGAGAAGAAATTCATT 189  
|||||  
|||  
Db 61 AGAGAGACTGTAAATAATGGCCACCCTTAGSAGATATAACCGAGAAGAAATTCATT 120

**Oy** 190 AGACAGACATCACACGCTGTGCCAGACTTGCCTTAACCAGAAGAAAGTAGTTGGCA 249  
|||||  
**Dd** 121 AGACAGACTTCACACAGCTGTGCCAGGCTTTGCATTAACCAAGAGACAGTAGTAGCTAACA 180

250 AGCACTGCTATGAGACACTGAGAAATTGTGTGGCCAAACAAACACTTGGCAATGSGACTCC 309  
 |||||  
 181 AGCACTGCTATGAGACACTGAAATAATTGTGTGGCCAAAGCACAACCTTGCATATGSGACTCC 240

Oy 310 AGTATGATTGTGCCCAACCAAGGAACCTCTCAGCAACCCTTGAAAAAGAAAAGAAGACTG 365  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 241 ACATCATTTGTCCCAAGCAGCGAAACTCTCAGCAACCTTGAGAAAGAAAAGAGACTG 300

QY 370 TGTGTCAATACTTTGACGACAGTGTCAAGTCAGATCAAGTGGAAATTTGTGAAACATCTT 422  
|||||  
|||||  
Db 301 TGTGTCAAGTATTTGACGACAGTGTCAAGTCTGTATCAAGTGGAAATTTGTGAAACACCTT 360

Qy 430 AATTCGCAATGTGCATTACCAACATGGGACATAAAGTGGTCTTAACCATATGTG 489  
||||| ||||| ||||| ||||| ||||| ||||| ||  
Db 361 AATTCGCAATGTGCATTACCAAGCATGGCACATCAACTCCTAAACCATATGTG 420

[illegible]

QY 550 CAGTCTACTGGAATGCCAAATCACTATGTGGTGACACTGTGTGCAGAAGATGGTAC 600  
||||| |||| |  
Db 481 CTGTCTACTTGGAGCCCAAGTCACGTGTGCTGCTGAGACCTCGTGTGCAAGGAATGGTAC 540

[illegible]

670 TCTCTGTGAGAGGCGCTGGCACAACGACAAAGGATGGGCGCATATTATTTAAAAACAAA / 2  
 QY  
 601 TCTCTGTGGCGAGCGCTGGCAGAGCGGCACAGGCTGGGGACAATACTTTATCAAAAACAAA 666  
 Db

Db 661 CCGTCGATGAGACGCTCCCACTCCTTTATAGAGCGCTTATTCCTAAAATCATTA 72

950 CAGACGCGCAGCTGGTAAATCCCAAGTCCTTTAATCATCTGTATGTAAGAATAATGCAAAAATA 90  
QY CAAGACATTGAGCAATAGACTTCTATTGGAGAAITGGGAAGCAATAGTITACATGAGAAI 84  
790 CAAGACATTGAGCAATAGACTTCTATTGGAGAAITGGGAAGCAATAGTITACATGAGAAI 84  
Db 721 CAAGACATTGAGCAATAGACTTCCAATTGGAGTGTTGGGCGCACATAGTTTACAGGAAATC 78

[illegible]

970 CCAATTTCTCAGCCGCAATAGGTTTACGCTTGTCTTCAGTATGATGAGAGAGATTC 10

[illegible]

1090 AACCGTTGATTCACCATTTGTGAAGCAGTCTCTCAGTTCGCTTCAATATAGCATGATG 11  
 Db 961 ATCCACAGCCTCTCAGACTCCACCGTCAGAGTGTGGAGTAAATCCAGGTGAGATGCTA 10  
 1030 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 11  
 1000 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 11

Db 1021 AACCATTTGATTACCACTGAAGCGTTCACACTGGCTTAATATGCATGANG 10

Ov 1150 GTACCTGTCCAAAGATCGTTCATGTGTATGGAGATATGCGTCCCAACTGCATT 12

[illegible]

```

Db 1141 ACCCTAGAGAGGCTGCTGGGAGACACGAGCTGCGCTCAATGTTGTACACTTTGATGAC 1200
OY 1270 AAGTACATGTTTCTGCATCTGGGGATAGACATATAAGGATGAGACCAAGTACTGT 1329
Db 1201 AAGTACATGTTTCTGCATCTGGGGATAGACATATAAGGATGAGACCAAGTACTGT 1260
OY 1330 GAATTTGTAAAGACCTTAAATATGAGACCAAGACGATTCCTGTTTGCAGTACAGGAG 1389
Db 1261 GAATTCGTAAAGACCTTAAATATGAGACCAAGGCTGGATGCTGTTTGCAGTACAGAGC 1320
OY 1390 AGGCTGTAGTACAGTGGCTATCTGACAAACACTATGATTAATGAGACATAGATGTGT 1449
Db 1321 AGGCTGTAGTACAGTGGCTATCTGACAAACACTATGATTAATGAGACATAGATGTGT 1380
OY 1450 GCATGTTTACGAGTGTAAAGGACCATGGAATGGGCTGATTCATTCATTAAC 1509
Db 1381 GCATGCTCGCACTGTTGAGAGGCGCATGAGAGTGTGATGCTGATTCATTAAC 1440
OY 1510 AAGAGATATGTCAGTCCGCTATGATGAGAAATTAAGTGTGGATCTTGCTGCT 1569
Db 1441 AAGAGATATGTCAGTCCGCTATGATGAGAAATTAAGTGTGGATCTTAATGCTGCT 1500
OY 1570 TTGGACCCCGCTGCTGCTGAGGAGACCTCTCTACGAGCCCTTGAGACATTCGGA 1629
Db 1501 TTGGACCCCGCTGCTGCTGAGGAGACCTCTCTGCGGACCTTGAGACATTCGGA 1560
OY 1630 AAGATTTTTCAGTACAGTGTGATTAATTCAGATTCATGATTCATGATGACCA 1689
Db 1561 AAGATTTTTCAGTACAGTGTGATTAATTCAGATTCATGATTCATGATGACCA 1620
OY 1690 ATTCATCTGTCGAGCTTCCCTAAATGATCCAGTCCGCAAGTGAACCCCGCTGCT 1749
Db 1621 ATTCATCTGTCGAGCTTCCCTAAATGATTCAGTCTGCTGCTGCTGCTGCTGCT 1680
OY 1750 TCTGACATACACTACATCCAGATTA 1779
Db 1681 TCTGACATACACTACATCCAGATTA 1710

RESULT 12
XELSBTRCP 1671 bp mRNA linear VRT 17-SEP-1993
LOCUS African clawed frog beta-trcp mRNA, complete cds.
DEFINITION M98268
ACCESSION M98268.1 GI:295542
VERSION beta-transducin repeats.
KEYWORDS Xenopus laevis (library: S. cerevisiae expression library of
SOURCE X.laevis oocytes) cDNA to mRNA.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodidae; Xenopus.
REFERENCE 1 (bases 1 to 1671)
AUTHORS Spevak,W., Keiper,B.D., Stratowa,C. and Castanon,M.J.
TITLE Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in
anaphase are rescued by Xenopus cDNAs encoding N-ras or a protein
with beta-transducin repeats
JOURNAL Mol. Cell. Biol. 13 (8), 4953-4966 (1993)
MEDLINE 9330289
PUBMED 8393141
FEATURES
source location/Qualifiers
1..1671 /organism="Xenopus laevis"
/db_xref="taxon:8355"
/cell_type="Oocyte"
/tissue_lib="S. cerevisiae expression library of X.laevis
oocytes"
1..1671 /gene="beta-trcp"
48..1604 /gene="beta-trcp"
/codon_start=1
CDS

```

```

/protein_id="AA02810.1"
/db_xref="GI:295543"
translation="MEFSCSLQPTASEREDCNREPPKLTTEKNTLROTKIANGT
SMIVPQRKLSANVEREKELCYKEFQMSSEDOVFVHLLSRKHCHGHNTYTK
PMLODFITLPAKRGDHLAENLSTYLDKSLASALYKREKRTPTSDMLKRLIER
MVRPDSIMRLARBRMGQYLFENKPPDGPFPNSYRALYPIIIOETIISNMG
RHSIORHCHSETSKVYCIQYDOKIVSGLRNTIRINDKNTLREKRVLMGTGSLV
CLODYDEVLIITGSSDSVRYMDVNTGEMNTLIHCEAVLIHFNMGMAVCSKDSI
AVMDMAVDITLRLVYVGHRAVNVNVEFDKTYIVASGDRITKVNTSTCEVRILN
GHRGIACIQYRDRLVYSSSDITILMFIIEGACLRLEHGEHLRCPTRPKRIVS
GAYDKIKYMDLVNADLPRAVATLTLRLTVEHSRVEFLQYIDEROYVSSSHDITLI
WDFLNDPGLA"
polyA_site 1671
BASE COUNT 495 a 341 c 424 g 411 t
ORIGIN
Query Match 47.3%; Score 1017.8; DB 5; Length 1671;
Best Local Similarity 81.8%; Pred. No. 1e-278;
Matches 1175; Conservative 0; Mismatches 262; Indels 0; Gaps 0;
OY 284 AAMCAAACTTGCCATGGCACTTCAGTATGATTTGCCCCAAGCAAGAACTCTCAG 343
Db 157 AGACAAACTTGCCATGGCACTTCAGTATGATTTGCCCCAAGCAAGAACTCTCAG 216
OY 344 CAAGCTATGAAAAAGAAAGAACTGTGTCAAAATCTTGAGCAGTGTGACAGTCA 403
Db 217 CAATTTGAGAAAGAAAGAAAGAACTGTGTCAAAATCTTGAGCAGTGTGACAGT 276
OY 404 ATCAAGTGAATTTGAGCACTTTATATCCCAATGTTGCTTACCAATGAGGACAA 463
Db 277 ATCAAGTGAATTTGAGCACTTTATATCCCAATGTTGCTTACCAATGAGGACAA 336
OY 464 TAACTGTATCTTAACTGTTGAGAGAGATTTAACTGCTGCGCAAGTCTGG 523
Db 337 TAAACACTTACCTTAAAGCAATGTAAAGAGATTTAACTGCTGCGCAAGTCTGG 396
OY 524 GATTTGATATGCTGAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTTGAGAG 583
Db 397 GATTTGATATGCTGAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTTGAGAG 456
OY 584 CTGAACTTGTGTCAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTTGAGAG 643
Db 457 CAGAACTGTATGTAAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTTGAGAG 516
OY 644 TTATGAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTTGAGAG 703
Db 517 TCATAGAGGAGATTTGAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTTGAGAG 576
OY 704 GGGGACAGTATTTATCAAAAACAACTCTGAGGGAATGCTCCCAACTCTTTT 763
Db 577 GGGGCAATATTTGTTTAAAAAACAACTCTGAGATGGGAAAAGCCCAAAATTTATCT 636
OY 764 ATAGACATTTATCTTAAATTTACAGACATTTGAGCAATTAATCTAATTTGAGAT 823
Db 637 ACAGACCGGTTTACCAAAAATTTATTAAGACATTTGAGCAATTTGAGCAATTTGAG 696
OY 824 GTGGAAGACATTTGAGCAATTTGAGCAATTTGAGCAATTTGAGCAATTTGAGCA 883
Db 697 GTGGAAGACATTTGAGCAATTTGAGCAATTTGAGCAATTTGAGCAATTTGAGCA 756
OY 884 GTTTACAGTATGATGATGAGAAATTAAGTAAAGGCTTTCGAGACAACTATCAATCT 943
Db 757 GTTCGACAGTATGATGATGAGAAATTAAGTAAAGGCTTTCGAGACAACTATCAATCT 816
OY 944 GGGATAAAAACATTTGAGCAATTTGAGCAATTTGAGCAATTTGAGCAATTTGAGCA 1003
Db 817 GGGATAAAAACATTTGAGCAATTTGAGCAATTTGAGCAATTTGAGCAATTTGAGCA 876
OY 1004 GTCTCCAGTATGATGAGAGATTTGATCAATGATTCATGATTTCCAGGCTCAGATGT 1063
Db 877 GTCTCCAGTATGATGAGAGATTTGATCAATGATTCATGATTTCCAGGCTCAGATGT 936

```

[illegible]

RESULT 14				
AB033280				
LOCUS	AB033280	2252 bp	mRNA	linear PRI 14-MAR-2000
DEFINITION	Homo sapiens BTGCP2 mRNA for F-box and WD-repeats protein			
ACCESSION	AB033280			
VERSION	AB033280.1 GI:7209810			
KEYWORDS	BTGCP2; F-box and WD-repeats protein beta-TrCP2 isoform B.			
SOURCE	Homo sapiens fetal lung cDNA to mRNA.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (stiles)			
REFERENCE	Kohle, N., Sagara, N., Kirikoshi, H., Takagi, A., Miwa, T., Hirai, M. and AUTHORS			

TITLE	Katoh, M.
JOURNAL	Molecular cloning and genomic structure of the betaTRCP2 gene on
MEDLINE	Chromosome 5q35.1
REFERENCE	Biochem. Biophys. Res. Commun. 269 (1), 103-109 (2000)
AUTHORS	2 (bases 1 to 2252)
JOURNAL	Katoh, M.
TITLE	Direct Submission
FEATURES	Submitted (05-OCT-1999) Masaru Katoh, National Cancer Center,
source	Genetics Division, Tsukiji 5-Chome, Chuo-Ku, Tokyo 104-0045, Japan
	(E-mail:mkatoh@nci.go.jp, Tel:81-3-3542-2511(ex.4402),
	Fax:81-3-3541-2685)
	Location/Qualifiers
	1..2252
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="5"
	/map="5q35.1"
	/lissue_type="Fetal Lung"
gene	1..2252
	/gene="BTRCP2"
	157..1746
	/gene="BTRCP2"
CDS	/codon_start=1
	/product="F-box and WD-repeats protein beta-TRCP2 isoform B"
	/protein_id="BAA92330.1"
	/db_xref="GI:7209811"
	/translation="MEPDSVIEDTTELMNTSYMEDNDPESKKKTLMTLOISNGSSS
	IVSRAPSGNTOKEKDCIKYEPDMSEDSOEPFVHLISRMCHROHGHINSLKPM
	ODPFTALPEGLDHIAENILISYLDASLCAELVCKEWMRYISBMWKILIERMVF
	THMLKIGSERGHQDYLFKRPFDGPNFSYLSVKPIIDLETIESMRCGRHNLQ
	RQCSSENSKYETCIQYDDEKTIISGLDNSKIKIMDTSECLKVLTVGHGYSYLQYI
	EAATVIGSSDSTVRVMDNNGEVLNLIHNEAVLIHFRSNGIYVSCSDRSIAWMD
	ASATYILSRVLVGHRAAVNVVDFDKYIVSASGDRITIKVMSITCEVPTLNGHKKR
	INCLQYRRLVYSGSDMTITLMDIECGCATLVIEGHEELVRCIFPDKRKIVSAYDQ
	KIKWMDLOALDPRAPASATICTRLTVLHSGRVRLQFDFEFOIISSHDDTILIMDFLN
	YVPSQNTNRSRSPRYTYSR"
BASE COUNT	598 a 501 c 558 g 595 t
ORIGIN	
Query Match	41.1%; Score 883.4; DB 9; Length 2252;
Best Local Similarity	74.8%; Pred. No. 2.3e-240;
Matches 1122; Conservative	0; Mismatches 371; Indels 6; Gaps 1;
QY . 298 AATGGCACTTCCAGTATGTGTGGCCCAAGAGCGAAACTCTCAGCAAGCTGTGAAG 357	
Db 271 AATGGAAATCATCTGTGATCTCTCCGAAAGAGGCCATCAGAAAGAACTATCAAAA 330	
QY 358 GAAAGAACTGTGTCAATACTTTGAGCAGTGTGCAGATCAGATCAAGTGAATT 417	
Db 331 GAAAGAACTGTGTATTAATATTTTACACAGTGTCTGAATCAGTCAAGTGAATT 390	
QY 418 GTGGAACATCTATATCCCAATGTGATATACCAATATGGGACATTAACTCGATCTT 477	
Db 391 GTGGAACATCTATTTTCAGATGTGTCAATATTCAGATGGAATATTAACCTTACT 450	
QY 478 AAACATATTTTCAGAGATTTTCAATACGTCCTGCCAGCTCGCGATTTGGATATATC 537	
Db 451 AAACCCATTTTCACGCGGACCTTATATACGCTTTACAGAGCAAGAGCTTAATACATA 510	
QY 538 GTCGAGAACATCTCTGATACCTGGATGGCAAAATCACTATGTGCTCTAATCTTGTGC 597	
Db 511 GCGAGAAACATTTCTTCTTACTCTGATGACCAAGGCTCTGTGTGCACCAACAGCTGTATGT 570	
QY 598 AAGGAATGTATCCGAGTACCTCGATGGATCGTGTGGAAGAGATTATCAGAGAAATG 657	
Db 571 AAGGAATGTGACGCGAGTATCTCAGAGGAATGCTTTGGAGAAGCTGATATGAAGAAATG 630	
QY 658 GTGACGACAGATTCCTGTGTGTGAGAGGCCCTGGCAGAAACGAAGAGATGGGAGCATTTTA 717	
Db 631 GTACGACATGATCCCTATGTGAAAGGACTTTGAAAGAAAGAGGGGTGGGATCAGTACTTG 690	



Db 473 AAGCCATGTTGACGGGAGCTTATTACGCTTTACGAGCAGGCTTATGATCAGATA 532  
QY 538 GCTGAGACATTTCTGTCATACCTGATAGCCAAATACATATGTCGTCACCTTGTGTC 597  
Db 533 GCAGAAAAACATTTCTTGTGTCATGATGATGATGATGATGATGATGATGATGATGAT 592  
QY 598 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 657  
Db 593 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 652  
QY 658 GTCAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 717  
Db 653 GTACGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 712  
QY 718 TTGAAAAACAAACCTCTGACGGGAAATGCTCTCCCACTCTTTTATAGACCTTTAT 777  
Db 713 TTTAAAAACAGACCCACAGATGCG-----CTCCAAATTCATTTTATAGTCATATATAC 766  
QY 778 CCTAAATTTATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837  
Db 767 CCAAGATTTATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 826  
QY 838 TTACAGAGATTTACCTGACGAGTGAAGCAAGAGATTTACTGTTTACAGATGAT 897  
Db 827 TTGAGAGATTTACCTGACGAGTGAAGCAAGAGATTTACTGTTTACAGATGAT 886  
QY 898 GATCAGAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 957  
Db 887 GATGAAAAAAATTTATCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 946  
QY 958 TTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017  
Db 947 CTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1006  
QY 1018 GAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077  
Db 1007 GAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1066  
QY 1078 GGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137  
Db 1067 GGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1126  
QY 1138 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197  
Db 1127 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1186  
QY 1198 CCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1257  
Db 1187 GCGACCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1246  
QY 1258 GACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1317  
Db 1247 GACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1306  
QY 1318 ACAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377  
Db 1307 ACAGACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1366  
QY 1378 CAGTACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1437  
Db 1367 CAGTACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1426  
QY 1438 ATAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1497  
Db 1427 ATAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1486  
QY 1498 CGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1557  
Db 1487 CGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1546  
QY 1558 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1617  
Db 1547 TTGCAAGCTGCTTTGACCTCGAGCCCGACAGACATTTGTTTGGCAGCATTTGGTG 1606

QY 1618 GAGCATTCGGAAGATTTTTCAGACTACAGTTGATGATGATGATGATGATGATGATGAT 1677  
Db 1607 GAACTTCTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1666  
QY 1678 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1737  
Db 1667 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1726  
QY 1738 CCCCCTTCCCTTCTGCAATACACCTACATCTCCAGATTAATTAACATACACTGAC 1796  
Db 1727 ACCGCTTCCCTCCAGAAATACATCTACATCTGATTAACAGTCTGCACTTTGAC 1785

Search completed: February 21, 2003, 23:50:02  
Job time : 5571 secs